

UNIVERSIDADE FEDERAL DO PARANÁ

LIDIA WATANABE REOLON

ESTRUTURA CRISTALOGRÁFICA DO FATOR DE INICIAÇÃO DA
TRADUÇÃO EIF4E5 DE *TRYPANOSOMA CRUZI* EM COMPLEXO COM O
CAP-4 DO MRNA E ESTUDOS DE INTERAÇÃO ENTRE AS HOMÓLOGAS
EIF4E DE *T. CRUZI* E ANÁLOGOS DO CAP.

CURITIBA

2019

LIDIA WATANABE REOLON

ESTRUTURA CRISTALOGRÁFICA DO FATOR DE INICIAÇÃO DA
TRADUÇÃO EIF4E5 DE *Trypanosoma cruzi* EM COMPLEXO COM O CAP-4
DO mRNA E ESTUDOS DE INTERAÇÃO ENTRE AS HOMÓLOGAS EIF4E DE
T. cruzi E ANÁLOGOS DO CAP.

Dissertação apresentada como requisito
parcial à obtenção do grau de Mestre
em Ciências -Bioquímica, no Curso de
Pós-Graduação em Ciências -
Bioquímica, Setor de Ciências
Biológicas, da Universidade Federal do
Paraná.

Orientadora: Prof.^a Dr.^a Beatriz Gomes
Guimarães

CURITIBA

2019

Universidade Federal do Paraná. Sistema de Bibliotecas.
Biblioteca de Ciências Biológicas.
(Giana Mara Seniski Silva – CRB/9 1406)

Reolon, Lidia Watanabe

Estrutura cristalográfica do fator de iniciação da tradução EIF4E5 de *Trypanosoma cruzi* em complexo com o cap-4 do mRNA e estudos de interação entre as homólogas EIF4E de *T. cruzi* e análogos do cap . / Lidia Watanabe Reolon. – Curitiba, 2019.

78 p.: il.

Orientador: Beatriz Gomes Guimarães

Dissertação (mestrado) - Universidade Federal do Paraná, Setor de Ciências Biológicas. Programa de Pós-Graduação em Ciências - Bioquímica.

1. Trypanosoma cruzi 2. Proteínas 3. Cristalografia I. Título II. Guimarães, Beatriz Gomes III. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de Pós-Graduação em Ciências-Bioquímica.

CDD (22. ed.) 616.9363



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO CIÊNCIAS
(BIOQUÍMICA) - 400010160/3P2

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em CIÊNCIAS (BIOQUÍMICA) da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **LIDIA WATANABE REOLON**, intitulada: **ESTRUTURA CRISTALOGRAFICA DO FATOR DE INICIAÇÃO DA TRADUÇÃO EIF4E5 DE *TRYPANOSOMA CRUZI* EM COMPLEXO COM O CAP-4 DO MRNA E ESTUDOS DE INTERAÇÃO ENTRE AS HOMÓLOGAS EIF4E DE *T. CRUZI* E ANÁLOGOS DO CAP.**, sob orientação da Profa. Dra. BEATRIZ GOMES GUIMARAES, após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de Mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 27 de Agosto de 2019.

BEATRIZ GOMES GUIMARAES
Presidente da Banca Examinadora

LUCIANO FERNANDES HUERGO
Avaliador Interno (UNIVERSIDADE FEDERAL DO
PARANÁ)

LETUSA ALBRECHT
Avaliador Externo (FUNDAÇÃO OSWALDO CRUZ)

AGRADECIMENTOS

Gostaria de agradecer primeiramente a minha família pelo apoio, em tudo, mas principalmente nos estudos. Agradeço a cada um, pelo convívio e por tudo aquilo que aprendemos juntos.

Agradeço a minha orientadora Dra. Beatriz Gomes Guimarães pela oportunidade e principalmente pela confiança em me aceitar neste projeto. Agradeço pela paciência em ensinar, e pelas dicas para a vida.

Agradeço aos colegas de laboratório, pelas risadas diárias, pelo apoio, pelas dicas e ajudas, pelos almoços e comemorações após expediente. Agradeço pela amizade sincera que formamos.

Agradeço a Pós-Graduação em Ciências – Bioquímica, ao Instituto Carlos Chagas FIOCRUZ-PR, a agência de fomento à pesquisa CNPq. Agradeço aos pesquisadores do Laboratório de Biologia Estrutural e Engenharia de Proteínas, ao Dr. Nilson Zanchin, e um agradecimento especial a Plataforma de Purificação de Purificação de Proteínas, em especial a Tatiana Assunção, a qual tem uma grande participação neste projeto.

Não tenha medo da perfeição.
Você nunca vai atingi-la.

(SALVADOR DALI)

RESUMO

A síntese de proteínas é um dos processos mais regulados pela célula. Os chamados fatores de iniciação da tradução (eIFs) participam da regulação desta etapa através de interações específicas necessárias à maquinaria de tradução. Os tripanossomatídeos possuem várias características moleculares únicas, um exemplo sendo sua maquinaria de síntese proteica. A iniciação da tradução é coordenada por 12 fatores de iniciação. A associação do mRNA ao ribossomo é mediada por um complexo multiproteico e dentre seus componentes está o fator de iniciação eIF4E, que interage especificamente com o chamado cap do mRNA. Uma das particularidades encontradas em tripanossomatídeos refere-se à estrutura do cap do mRNA, denominado cap-4, que possui, além do 7-metil-GTP dos eucariotos, metilações adicionais nos quatro primeiros nucleotídeos do *spliced leader*. Outra particularidade dos tripanossomatídeos é a presença de seis homólogos eIF4E (chamadas EIF4E1 a 6), que podem estar envolvidas em diferentes complexos pré-iniciação ditos eIF4F-like. Assim, o projeto tem como objetivo estudar os 6 diferentes homólogos de eIF4E de *Trypanosoma cruzi* e suas interações com análogos do cap. Ensaio de interação entre EIF4E e os análogos do cap foram realizados por termoforese em microescala, além de ensaios de cristalização, que levaram a determinação da estrutura cristalográfica de TcEIF4E5. A estrutura cristalográfica do complexo TcEIF4E5 com os análogos do cap do mRNA revelou detalhes únicos do mecanismo de interação, a estrutura cristalográfica apresentou pela primeira vez a estrutura tridimensional de EIF4E em complexo com cap-4. As análises de interação das construções de EIF4E por termoforese em microescala indicam regiões envolvidas na interação e sugerem papéis diferenciados entre cada homólogo do fator eIF4E em *T. cruzi*.

Palavras chave: Fatores de Iniciação da Tradução, eIF4E, cap-4, *Trypanosoma cruzi*, Estrutura Cristalográfica de TcEIF4E5.

ABSTRACT

Protein synthesis is one of the most regulated processes in the cell. The so-called translation initiation factors (eIFs) participate in the regulation of this stage through specific interactions, necessary to the translation machinery. Trypanosomatids have several unique molecular characteristics, including their protein synthesis machinery. Translation initiation is coordinated by 12 initiation factors. The association of the mRNA with the ribosome is mediated by a multiprotein complex and among its components is the eIF4E initiation factor, which interacts specifically with the mRNA cap. One of the particularities found in trypanosomatids refers to the mRNA cap structure, called cap-4, which has in addition to the eukaryote 7-methyl-guanosine, additional methylations on the first four nucleotides of the spliced leader. Another particular feature of trypanosomatids is the presence of six eIF4E homologs (called EIF4E1 to 6), which may be involved in different pre-initiation complexes called eIF4F-like. This project aims to study the 6 different eIF4E homologues of *Trypanosoma cruzi* and their interactions with cap analogs. Interaction analysis between EIF4E and cap analogs were performed by microscale thermophoresis, as well as crystallization assays, which led to the determination of the TcEIF4E5 crystallographic structure. The crystal structures of the TcEIF4E5 in complex with mRNA cap analogues provided unprecedented details on the binding mechanism and revealed for the first time the three-dimensional structure of a trypanosomatid eIF4E in complex with the mRNA cap-4. Interaction analyzes of EIF4E constructs by microscale thermophoresis indicate regions involved in the interaction and suggest different roles between each eIF4E factor homologue in *T. cruzi*.

Keywords: Translation Initiation Factors, eIF4E, cap-4, *Trypanosoma cruzi*, TcEIF4E5 Crystallographic Structure.

LISTA DE FIGURAS

Figura 1 – Representação da estrutura química do cap do mRNA.	15
Figura 2 - Representação esquemática do processo de iniciação da tradução em eucariotos..	17
Figura 3 - Estrutura do cap-4 do mRNA de tripanossomatídeos.	19
Figura 4 - Representação da estrutura tridimensional de eIF4E humana e complexo com m ⁷ GTP e peptídeo de eIF4G1 (PDB 2w97).	22
Figura 5 - Detalhes da interação eIF4E - cap.....	23
Figura 6 - Estrutura tridimensional de LmEIF4E1 (PDB: 5WB5).	24
Figura 7 - Representação esquemáticas dos seis fatores EIF4E de <i>T. cruzi</i>	27
Figura 8 – Análise dos testes de expressão TcEIF4E2ΔN (28 kDa) em SDS-PAGE.....	38
Figura 9 – Análise dos testes de expressão das construções TcEIF4E5 (24 kDa) e TcEIF4E5ΔC (23 kDa) em SDS-PAGE.....	38
Figura 10 – Análise dos resultados dos testes de purificação das construções TcEIF4E1ΔN e TcEIF4E1ΔINS por SDS-PAGE.....	39
Figura 11 - Análise dos resultados dos testes de purificação das construções TcEIF4E2ΔNC e TcEIF4E2ΔS1-S2 por SDS-PAGE.....	40
Figura 12 – Análise dos resultados dos testes de purificação das construções TcEIF4E3ΔN e TcEIF4E4ΔN por SDS-PAGE.	40
Figura 13 - Análise do resultado do teste de purificação da construção TcEIF4E6 por SDS- PAGE.....	41
Figura 14 - Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1 por afinidade a níquel.	42
Figura 15 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1 por exclusão de tamanho em coluna Superdex 75 16/600.	42
Figura 16 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔN por afinidade a níquel.	43
Figura 17 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔN por troca iônica.	44
Figura 18 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔN por exclusão de tamanho em coluna Superdex 75 16/600.	44
Figura 19 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔS4-H2 por afinidade a níquel.....	45
Figura 20 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔS4-H2 por troca iônica.....	45
Figura 21 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔS4-H2 por exclusão de tamanho, em coluna Superdex 75 16/600.	46
Figura 22 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E2ΔN por afinidade a níquel.	47

Figura 23 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E2ΔN por exclusão de tamanho em coluna Superdex 75 16/600.	47
Figura 24 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E2ΔS1-S2 por afinidade a níquel.	48
Figura 25 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E2ΔS1-S2 por exclusão de tamanho em coluna Superdex 75 16/600.	48
Figura 26 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E3ΔN.	49
Figura 27 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E3ΔN por exclusão de tamanho em coluna Superdex 75 Increase 10/300.	49
Figura 28 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E3ΔN.	50
Figura 29 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E4ΔN.	51
Figura 30 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E4ΔN por exclusão de tamanho, em coluna Superdex 75 Increase 10/300.	51
Figura 31 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E5 por afinidade a níquel.	52
Figura 32 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E5 por exclusão de tamanho, em coluna Superdex 75 16/600.	52
Figura 33 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E5ΔC por afinidade a níquel.	53
Figura 34 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E5ΔC por exclusão de tamanho, em coluna Superdex 75 16/600.	54
Figura 35 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E6.	54
Figura 36 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E6 por exclusão de tamanho, em coluna Superdex 75 Increase 10/300.	55
Figura 37 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E6.	55
Figura 38 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E1.	57
Figura 39 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E2.	58
Figura 40 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E5.	60
Figura 41 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E3 e TcEIF4E6.	61
Figura 42 - Cristais obtidos de TcEIF4E2ΔN.	63
Figura 43 - Cristais obtidos de TcEIF4E1ΔN.	64
Figura 44 - Cristais obtidos em testes preliminares de TcEIF4E5.	64
Figura 45 - Cristais obtidos de TcEIF4E5 marcados com selênio-metionina.	65
Figura 46 - Cristais obtidos de TcEIF4E5 com análogos do cap.	65

LISTA DE TABELAS

Tabela 1 - Interação entre os homólogos dos fatores EIF4E e EIF4G de tripanossomatídeos descritos na literatura.	21
Tabela 2 - Constantes de dissociação dos fatores eIF4E com análogos do cap.	21
Tabela 3 - Descrição das construções dos fatores EIF4E de <i>T. cruzi</i> utilizados neste trabalho.	28
Tabela 4 - Composição dos Tampões A de cromatografia de afinidade.	30
Tabela 5 - Composição dos Tampões B utilizadas nas cromatografias de afinidade.....	30
Tabela 6 - Tampões utilizados na purificação por gravidade.	32
Tabela 7 - Condições de cristalização utilizados no primeiro refinamento de TcEIF4E1ΔN.	35
Tabela 8 - Condições de cristalização utilizados no segundo refinamento de TcEIF4E1ΔN, com a concentração de 4,8 mg/mL.	35
Tabela 9 – Condições de cristalização utilizados no segundo refinamento de TcEIF4E1ΔN, com a concentração de 4 mg/mL.	36
Tabela 10 - Valores de constante de dissociação obtidos para as construções TcEIF4E1, TcEIF4E1ΔN e TcEIF4E1ΔS4-H2 testados com os análogos m ⁷ GTP e cap-4 por termoforese em microescala.....	57
Tabela 11 - Valores da constante de dissociação (K _d) para interação do fator EIF4E de tripanossomatídeos com análogos do cap.....	62

LISTAS DE ABREVIATURAS E SIGLAS

D.O.600: densidade ótica medida a 600 nm

DTT: 1,4-Ditiotreitol

eIF: fator de iniciação em eucariotos, do inglês eukaryotic initiation fator

FPLC: cromatografia líquida rápida de proteínas, do inglês fast protein liquid chromatography

IPTG: Isopropil- β -tiogalactopiranosídeo

Kd: constante de dissociação

m⁷GTP: 7-metil guanosina trifosfato

PDB: Protein Data Bank

PEG: polietilenoglicol

SDS: dodecil sulfato de sódio

SDS-PAGE: eletroforese em gel de poliacrilamida (do inglês polyacrylamide gel electrophoresis) em presença de detergente SDS.

VC: volume de coluna

SUMÁRIO

1	Introdução.....	14
1.1	Tradução em eucariotos	14
1.1.1	Maturação do RNA mensageiro.....	15
1.1.2	Maquinaria de tradução	16
1.2	Tradução em tripanossomatídeos	18
1.2.1	A estrutura cap-4 do mRNA.....	19
1.2.2	Complexo eIF4F em tripanossomatídeos	20
1.3	Características estruturais do fator de iniciação da tradução eIF4E	22
2	Justificativa	25
3	Objetivos.....	26
3.1	Objetivo geral.....	26
3.2	Objetivos específicos	26
4	Materiais e métodos.....	27
4.1	Desenho das construções para expressão dos fatores EIF4E de <i>T. cruzi</i>	27
4.2	Testes de expressão.....	28
4.3	Expressão das construções de EIF4E de <i>T. cruzi</i>	29
4.4	Purificação das proteínas de interesse	30
4.4.1	Purificação em sistema automatizado	30
4.4.2	Purificação por gravidade	32
4.5	Ensaio de interação por termoforese em micro-escala.....	33
4.6	Ensaio de cristalização	34
4.6.1	TcEIF4E2ΔN	34
4.6.2	TcEIF4E1ΔN e TcEIF4E1ΔS4-H2	34
4.6.3	TcEIF4E5	36
5	Resultados e discussão	38
5.1	Testes de expressão.....	38
5.1.1	Expressão em cepa <i>E. coli</i> BL21(DE3) STAR	38
5.1.2	Expressão e teste de purificação em cepa <i>E. coli</i> BL21 ΔSlyD pRARE2	39
5.2	Expressão e purificação.....	41
5.2.1	TcEIF4E1, TcEIF4E1ΔN e TcEIF4E1ΔS4-H2.....	41

5.2.2	TcEIF4E2 Δ N e TcEIF4E2 Δ S1-S2.....	46
5.2.3	TcEIF4E3 Δ N	49
5.2.4	TcEIF4E4 Δ N	50
5.2.5	TcEIF4E5 e TcEIF4E5 Δ C	52
5.2.6	TcEIF4E6	54
5.3	Termoforese em microescala	56
5.3.1	TcEIF4E1, TcEIF4E1 Δ N e TcEIF4E1 Δ S4-H2.....	56
5.3.2	TcEIF4E2 Δ N e TcEIF4E2 Δ S1-S2.....	58
5.3.3	TcEIF4E5 e TcEIF4E5 Δ C	59
5.3.4	TcEIF4E3 Δ N e TcEIF4E6.....	60
5.4	Ensaio de cristalização.....	63
5.4.1	TcEIF4E2 Δ N	63
5.4.2	TcEIF4E1 Δ N	63
5.4.3	TcEIF4E5	64
5.5	Determinação e análise da estrutura cristalografica de TcEIF4E5	66
6	Considerações finais.....	68
7	Referências.....	70
ANEXO A – Protocolo de purificação por gravidade.....		75
ANEXO B – Protocolo de expressão em meio mínimo M9, e purificação de TcEIF4E5-SeMet.....		76
ANEXO C - Artigo Crystal structure of the <i>Trypanosoma cruzi</i> EIF4E5 translation factor homologue in complex with mRNA cap-4		78

1 INTRODUÇÃO

Os tripanossomatídeos formam um grupo de protozoários que compreende várias espécies patogênicas dos gêneros *Leishmania* e *Trypanosoma*, responsáveis por doenças humanas graves e negligenciadas. Os tripanosomas são conhecidos desde 1841, destaca-se a descoberta de Carlos Chagas em colaboração com Oswaldo Cruz, da espécie *Trypanosoma cruzi* em 1909. A doença de Chagas é ainda muito recorrente nos países da América, sendo este um dos motivos de muitos estudos sobre o *T. cruzi* (JAHNE et al. 2016; LANA e MACHADO, 2010). Além disso, os tripanossomatídeos apresentam aspectos únicos em relação ao seu ciclo de vida e mecanismos moleculares e algumas destas singularidades estão relacionadas à sua maquinaria de síntese proteica.

A síntese dos RNA mensageiros (mRNA) em tripanossomatídeos ocorre de forma policistrônica, onde um único transcrito contém vários genes. Acredita-se que por este motivo o controle pós-transcricional seja tão regulado, principalmente no nível traducional. Outra diferença na síntese do mRNA em tripanossomatídeos é a estrutura do cap do mRNA, denominado cap-4. A síntese de proteínas, não apenas em tripanossomatídeos, mas também em outros eucariotos é um dos processos mais regulados pela célula, e envolve grande gasto de energia. Os chamados fatores de iniciação da tradução (eIFs) participam da regulação desta etapa através de interações transientes, mas específicas, necessárias à maquinaria de tradução. Os tripanossomatídeos neste ponto se destacam pela presença de vários homólogos dos fatores de iniciação da tradução (ALBERTS et al. 2010; HINNEBUSCH, 2014; ZINOVIEV e SHAPIRA, 2012).

1.1 TRADUÇÃO EM EUCARIOTOS

Em geral, o controle da expressão gênica ocorre no processo de transcrição. Por meio de interações específicas ao DNA pode-se favorecer ou não a transcrição, sendo essas interações específicas a partir de sinalização

celular, devido as condições em seu meio (ALBERTS et al. 2010; ZAHA, FERREIRA e PASSAGLIA, 2012).

Em eucariotos o processo de tradução, através do qual ocorre a síntese de proteínas, pode ser dividido em três etapas principais, a iniciação, o alongamento da cadeia peptídica e o término. A molécula de mRNA é onde se encontra a informação para que o ribossomo, uma maquinaria catalítica, transcreva as informações em polipeptídios. O processo da tradução é complexo e muito bem controlado por diversas proteínas e fatores, de modo que a célula tenha uma rápida resposta a sinais internos e externos ao seu meio (ALBERTS et al. 2010; ZAHA, FERREIRA e PASSAGLIA, 2012).

1.1.1 Maturação do RNA mensageiro

Após a síntese do mRNA no núcleo da célula é necessária à sua exportação ao meio citoplasmático, onde se encontram as demais moléculas que realizam a síntese de proteínas (LEHNINGER, NELSON e COX, 2011).

O processo de maturação do mRNA envolve a adição de um cap à extremidade 5' e uma cauda rica em adenosinas (poli-A) à sua extremidade 3'. A cauda poli-A consiste em cerca de 80 a 250 resíduos de Adenina. Durante a maturação do mRNA ocorre também a excisão de íntrons e a ligação dos éxons (AYLETT e BAN, 2017).

A estrutura do cap do mRNA, ilustrado na Figura 1, consiste em uma 7-metil guanosina ou m⁷GpppN (Onde N é uma das quatro bases nitrogenadas). A adição do cap ao mRNA recém-sintetizado dá-se por meio de uma ligação 5'-5' trifosfato no primeiro nucleotídeo transcrito, e onde a guanosina é metilada na posição N-7 (LIU et al., 2009; ALBERTS et al. 2010).

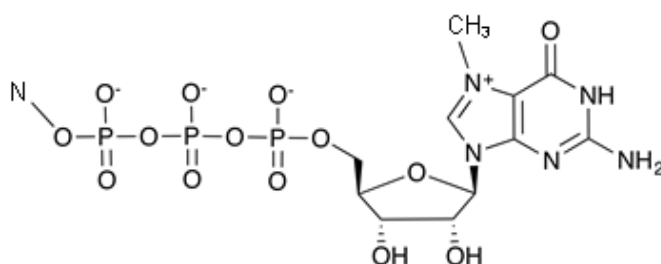


Figura 1 – Representação da estrutura química do cap do mRNA.

A estrutura do cap o qual é ligado ao RNA mensageiro antes da sua exportação ao citoplasma tem função de proteção do mRNA da ação de exonucleases, identificador para o recrutamento de fatores do splicing do pré-mRNA e do transporte dos mRNAs maduros através da membrana nuclear, também promovem o processo da tradução e participam do complexo de iniciação da tradução, tendo um importante papel na expressão dos genes (ALBERTS et al. 2010). Em eucariotos é comum a presença do 5'-cap nos mRNAs que se encontram no citoplasma, porém não ocorrem em mRNAs de cloroplastos e mitocôndrias (ZAHA, FERREIRA e PASSAGLIA, 2012).

1.1.2 Maquinaria de tradução

O processo de tradução possui vários pontos de regulação, ao menos 12 fatores de iniciação coordenam o início da tradução, atuando na associação do ribossomo ao RNA mensageiro e regulando os mecanismos para reconhecimento do códon de iniciação (AYLETT e BAN, 2017).

Para que se inicie o processo de tradução é necessária a coordenação de várias interações proteicas entre os fatores de iniciação, (eIF, do inglês eukariotic initiation factors), formando complexos que estão envolvidos nesta etapa. Entre os complexos formados dois se destacam, o eIF4F e o complexo de pré-iniciação (PIC, do inglês Pré-initiation complex) (AYLETT e BAN, 2017; HINNEBUSCH, 2014).

eIF4F é composto por eIF4E (proteína que se liga diretamente ao cap do mRNA), eIF4G (proteína de ancoragem) e por eIF4A (uma helicase responsável por desfazer as estruturas secundárias do mRNA) (Figura 2, (2) e (4)). O complexo eIF4F interage também com a proteína PABP (Poli-A binding protein) a qual se liga a extremidade 3' do mRNA, via cauda poli-A. A interação de eIF4F e PABP mantem o mRNA em uma forma circular (AITKEN; LORSCH, 2012; LEHNINGER, NELSON e COX, 2011).

O complexo de pré-iniciação PIC é composto por Met-tRNA, pelos fatores eIF1A, eIF2 (responsável por catalisar a adição de metionina ao tRNA iniciador, por meio de quebra de GTP), eIF5 e eIF3 (responsável por recrutar a subunidade menor (40S) do ribossomo) (Figura 2, (1) e (3)). PIC associa-se ao mRNA por

interações com o complexo eIF4F, mais especificamente pela interação entre eIF3 de PIC e eIF4G do complexo eIF4F, o qual então permite a varredura do mRNA até o encontro do códon de iniciação (ALBERTS et al. 2010; ZAHA, FERREIRA e PASSAGLIA, 2012; HINNEBUSCH, 2014).

Havendo o reconhecimento do *start códon* pelo PIC a subunidade maior do ribossomo (60S) é recrutada, tornando-o completo. Iniciando assim a progressão da elongação da tradução, etapa que é irreversível. A partir deste momento os fatores eIF5B e eIF6 atuam liberando a maioria dos fatores de iniciação para reciclagem (AYLETT e BAN, 2017; ZAHA, FERREIRA e PASSAGLIA, 2012).

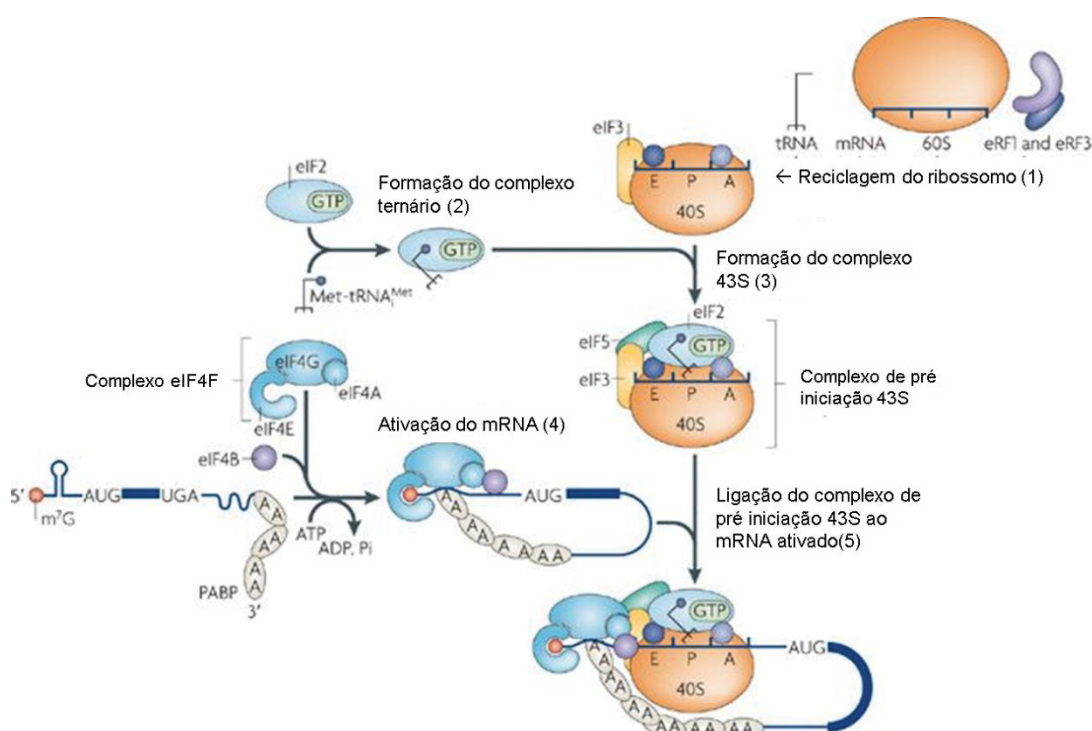


Figura 2 - Representação esquemática do processo de iniciação da tradução em eucariotos. 1) Reciclagem do Ribossomo. 2) Formação do complexo ternário (composto por eIF2-GTP e Met-tRNA). 3) Formação do complexo 43S. 4) Ativação do mRNA. 5) Ligação do complexo de pré iniciação 43S ao mRNA ativado. Fonte: Jackson et al., 2010.

Para o processo de início da tradução, no entanto, é necessário o reconhecimento do cap do mRNA pelo fator eIF4E para a correta seleção dos mRNAs que devem ser traduzidos. A desregulação desta função do fator 4E, bem como de outras proteínas que se ligam ao cap, está associada a problemas fisiológicos como a proliferação de células cancerígenas (CAJAL et al., 2018).

Os fatores eIF4E em mamíferos possuem mais de uma homóloga e normalmente são classificados em três classes (JOSHI et al., 2005).

A classe I corresponde ao chamado eIF4E canônico, denominado eIF4E1, que desempenha o papel descrito anteriormente no processo de iniciação da tradução (JOSHI, 2005; TOMOO, 2005). A classe II engloba proteínas homólogas denominadas eIF4E2 ou 4EHP (4E homologous protein) que apresentam substituições em resíduos de triptofano responsáveis pela interação com o cap. Tais resíduos podem ser substituídos por tirosina, fenilalanina ou leucina, o que afeta a afinidade destas proteínas pelo cap do mRNA (ROSETTANI, 2007). A classe III, denominada eIF4E3 é caracterizada pelos homólogos que possuem um resíduo não aromático cisteína (C52 em eIF4E3) em substituição ao triptofano responsável na interação com o cap. eIF4E3 humano também foi descrito como inibidor da tradução (OSBORNE, 2013).

O parceiro de interação de eIF4E, eIF4G, também possui mais de um homólogo caracterizado. Em mamíferos foram descritos três, os quais contêm domínios conservados para interação com o complexo PIC e com a proteína eIF4A (AITKEN; LORSCH, 2012; KLANN; DEVER, 2004).

1.2 TRADUÇÃO EM TRIPANOSSOMATÍDEOS

O ciclo de vida do *T. cruzi* varia entre os hospedeiros, e apresenta três morfologias principais, tripomastigota (forma infectiva), amastigota (forma replicativa) e epimastigota (forma não infectiva encontrado no inseto vetor) (LANA e MACHADO, 2010).

Devido a alteração de sua morfologia entre os hospedeiros, as condições de vida como parasitas e por serem organismos unicelulares os tripanossomatídeos desenvolveram vários mecanismos adaptativos, possuindo modos característicos de expressarem seus genes. A transcrição dos genes codificadores de proteínas em tripanossomatídeos ocorrem policistronicamente, e são processados por *trans*-splicing e poliadenilação (ZINOVIEV e SHAPIRA, 2012).

1.2.1 A estrutura cap-4 do mRNA

A maturação do mRNA também faz parte da regulação da expressão genica em tripanossomatídeos (GOMEZ et al, 2010). Os mRNAs policistrônicos são processados por trans-splicing.

Trans-splicing é um processo que envolve a adição na região 5' de cada unidade do precursor policistrônico de um éxon conservado (de 39 nucleotídeos) derivado de um *spliced leader* (SL) RNA. O SL RNA antes de ser adicionado aos genes independentes do pré-mRNA recebe metilações na extremidade 5' formando uma estrutura do cap diferenciado dos demais eucariotos, denominado cap-4 (Figura 3) (LEHNINGER, NELSON e COX, 2011).

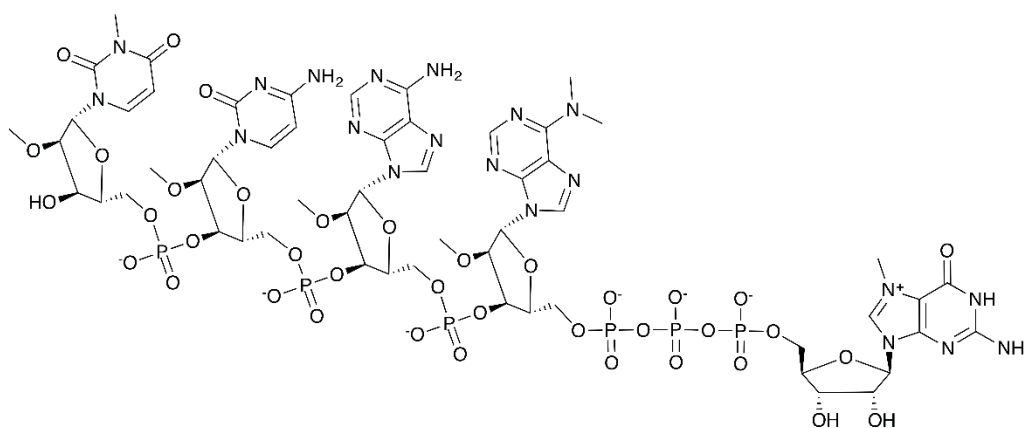


Figura 3 - Estrutura do cap-4 do mRNA de tripanossomatídeos.

O cap-4 gerado pelo processo de trans-splicing possui 2'-O ribose metiladas nos quatro primeiros nucleotídeos do mRNA e também 2 metilações na primeira adenosina e uma na quarta uridina. Essas metilações não ocorrem em outros grupos de eucariotos (BANGS et al., 1992; ZINOVIEV e SHAPIRA, 2012).

Devido às metilações no cap-4 é necessário um reconhecimento diferenciado do fator eIF4E para o recrutamento dos fatores da maquinaria de tradução. Os vários homólogos encontrados do fator eIF4E indicam interações com afinidades diferentes ao cap (LIU et al., 2009; TOMOO et al. 2005).

1.2.2 Complexo eIF4F em tripanossomatídeos

Em tripanossomatídeos são conhecidos 6 homólogos de eIF4E para *Trypanosoma brucei*, *Trypanosoma cruzi* e para *Leishmania major*, e 5 homólogos de eIF4G foram identificados. eIF4E e eIF4G interagem na formação do complexo eIF4F, sendo que diferentes homólogos apresentam interações específicas (FREIRE et al. 2011a; ZINOVIEV e SHAPHIRA, 2012).

Os 6 homólogos do fator eIF4E tem sido classificados em 3 grupos, devido às suas propriedades moleculares (FREIRE et al. 2017). EIF4E1 e EIF4E2 do grupo I são expressos em baixos níveis e estão localizados tanto no núcleo quanto no citoplasma. São os mais similares com os eIF4Es de mamíferos, fungos e plantas. Foi caracterizada uma fraca interação dos fatores EIF4E1 e EIF4E2 de *T. brucei* com os fatores EIF4G3 e EIF4G4 (DHALLIA et al. 2005; FREIRE et al. 2011). E em *Leishmania major* EIF4E1 interage com o homólogo EIF4G3 (YOFFE et al., 2009), as interações específicas descritas entre EIF4E e EIF4G em tripanossomatídeos estão relacionadas na Tabela 1.

No grupo II, EIF4E3 e EIF4E4 são abundantes no citoplasma e interagem com homólogos de EIF4G específicos. Possuem um N terminal extenso envolvido na interação como homólogos de PABP, essa mesma longa extensão N-terminal também sendo relacionada à regulação da tradução através de processos de fosforilação (FREIRE et al. 2011, ZINOVIEV et al. 2011). Os pares de interação descritos são LeishIF4E4 - LeishIF4EG3, sendo este complexo o melhor candidato para a via de início da tradução. Ambos fatores 4E3 de *T. brucei* e *Leishmania* interagem com o fator 4G4 (DE MELO NETO, 2015; YOFFE et al. 2006, 2009).

No grupo III, EIF4E5 e EIF4E6 são estruturalmente menores, são encontrados no citoplasma e interagem com homólogos de EIF4G específicos (FREIRE et al. 2014a, b). Já foram descritas as interações de TbEIF4E5 com EIF4G1 e EIF4G2, enquanto TbEIF4E6 interage exclusivamente com EIF4G5. (ZINOVIEV e SHAPHIRA, 2012; FREIRE, et al. 2011; FREIRE, et al. 2014b).

Organismo	Fator	Resultado	Referência
<i>T. brucei</i>	EIF4E1	Interação com eIF4G3 e eIF4G4	Freire et al., 2011
<i>T. brucei</i>	EIF4E2	Interação com eIF4G3 e eIF4G4	Freire et al., 2011
<i>T. brucei</i>	EIF4E3	Interação com eIF4G4 e eIF4G3	Freire et al., 2011
<i>T. brucei</i>	EIF4E4	Interação com eIF4G3	Freire et al., 2011
<i>T. brucei</i>	EIF4E5	Interação com eIF4G2 e eIF4G1	Freire et al., 2014b
<i>T. brucei</i>	EIF4E6	Interação com eIF4G5	Freire et al., 2014c
<i>L. major</i>	EIF4E1	Interação com eIF4G3	Yoffe et al., 2009
<i>L. major</i>	EIF4E4	Interação com eIF4G3	Yoffe et al., 2009

Tabela 1 - Interação entre os homólogos dos fatores EIF4E e EIF4G de tripanossomatídeos descritos na literatura.

Ensaio biofísicos com quatro homólogos de eIF4E utilizando cap-4 e análogos quimicamente sintetizados mostraram uma variedade na afinidade de ligação aos análogos do cap (Tabela 2).

Organismo	Fator	Kd GTP (μM)	Kd m ⁷ GTP (μM)	Kd cap4 (μM)	Referência
<i>L. major</i>	LmIF4E1	300	5,98	4,35	Método: Time synchronized fluorescence titration Yoffe et al., 2006
<i>L. major</i>	LmIF4E2	300	27	3,57	
<i>L. major</i>	LmIF4E3	100	3,12	21,74	
<i>L. major</i>	LmIF4E4	100	1,25	1,30	
<i>T. brucei</i>	TbIF4E5	27	1,54	1,82	Método: Time synchronized fluorescence titration Freire et al., 2014a
<i>T. brucei</i>	TbIF4E6	-	6,25	6,25	Método: Time synchronized fluorescence titration Freire et al., 2014b
<i>H. sapiens</i>	HsIF4E1	-	0,0085	-	Método: Time synchronized fluorescence titration Niedzwiecka et al., 2002
<i>H. sapiens</i>	HsIF4E2	58	2,38	-	Método: Isothermal titration calorimetry Rosettani et al., 2007
<i>M. musculus</i>	MmIF4E1	31	0,0082	0,038	Método: Time synchronized fluorescence titration Yoffe et al., 2006
<i>M. musculus</i>	MmIF4E3	-	1,8	-	Método: Isothermal titration calorimetry Osborne et al 2013

Tabela 2 - Constantes de dissociação dos fatores eIF4E com análogos do cap.

Os vários homólogos encontrados para eIF4E em tripanossomatídeos indicam adaptações evolutivas de modo a se ajustar a conformação necessária

para a ligação com o cap-4 (LIU et al., 2009). Os resultados descritos na literatura evidenciam uma afinidade menor dos fatores EIF4E de tripanossomatídeos ao cap, quando comparados aos homólogos dos demais eucariotos.

1.3 CARACTERÍSTICAS ESTRUTURAIS DO FATOR DE INICIAÇÃO DA TRADUÇÃO EIF4E

A primeira estrutura cristalográfica de eIF4E determinada revelou a conformação da proteína em um formato de C (MARCOTRIGIANO, 1997) (Figura 4), possuindo extremidades flexíveis, sendo a extremidade C-terminal mais conservada que a porção N-terminal. A estrutura do domínio da família eIF4E é conservada e se caracteriza por 7 fitas- β antiparalelas e 3 ou mais α -hélices.

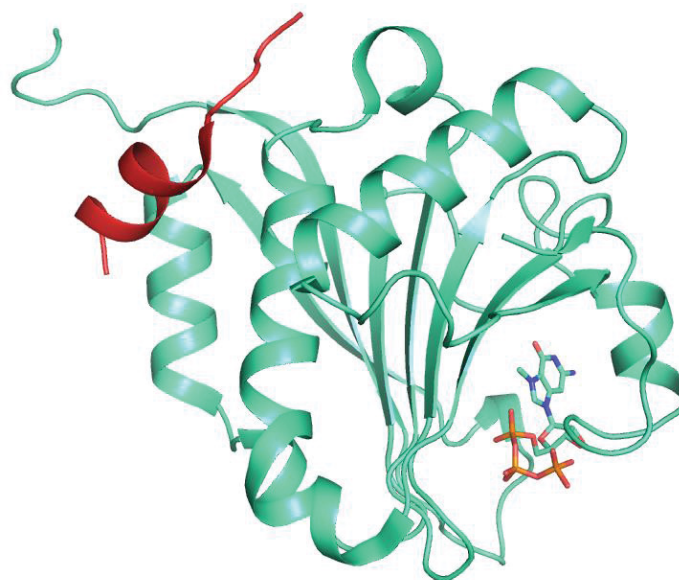
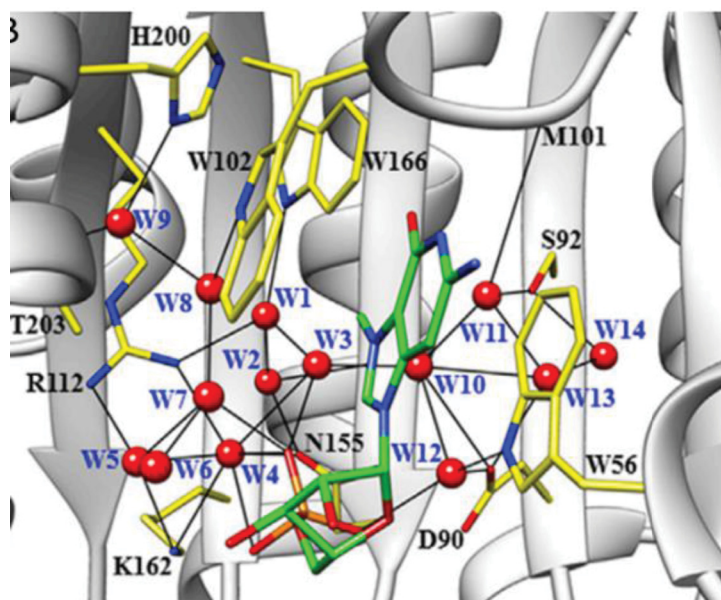


Figura 4 - Representação da estrutura tridimensional de eIF4E humana (em verde) e complexo com m⁷GTP (*sticks* em verde) e peptídeo de eIF4G1 (em vermelho) (PDB 2w97).

O sítio de ligação ao cap se encontra no lado côncavo da proteína, onde se encontram resíduos de triptofano essenciais (posição W56, W102 e W166 em eIF4E1 humana) para a ligação com o mRNA, estes resíduos formam uma ligação do tipo sanduiche com a guanina do cap. O metil na posição 7 do cap participa de uma interação específica com um glutamato muito conservado (E103 em eIF4E1 humana). Além das interações envolvendo a guanina há também interações entre os resíduos básicos de arginina (posições R112 e R157



A estrutura cristalográfica determinada por Tomoo e colaboradores (2002) mostrou a importância da região C terminal na interação com os análogos do cap do mRNA. Os resíduos indicados como responsáveis pela interação da região C-terminal de eIF4E1 humano são T203, T205 e S209, destaca-se também o resíduo H200 pelo direcionamento da posição de W102 o qual participa no *stacking* da interação com o cap.

Estruturas cristalográficas de eIF4E em complexo com peptídeos de eIF4G e de 4E-BP (4E binding protein) revelam a região de interação entre os parceiros. A região N-terminal de eIF4G possui a sequência consenso YXXXXLΦ (onde X: qualquer aminoácido e Φ: aminoácido hidrofóbico), que forma uma hélice canônica a qual interage com a proteína eIF4E pelo motivo VXXFW (onde X: qualquer aminoácido), situado no lado oposto ao sítio de ligação ao cap (Figura 4). As proteínas 4E-BP competem com o fator 4G na interação com eIF4E, tal interação é descrita como reguladora do processo de tradução (SAFAEE et al., 2012).

Recentemente, foi descrita a primeira estrutura tridimensional de uma homóloga de eIF4E de tripanossomatídeos, a estrutura de EIF4E1 de *Leishmania major* (PDB: 5WB5) (MELEPPATTU, 2018). A estrutura foi determinada na sua forma apo (na ausência de análogos do cap) e está representada na Figura 6.

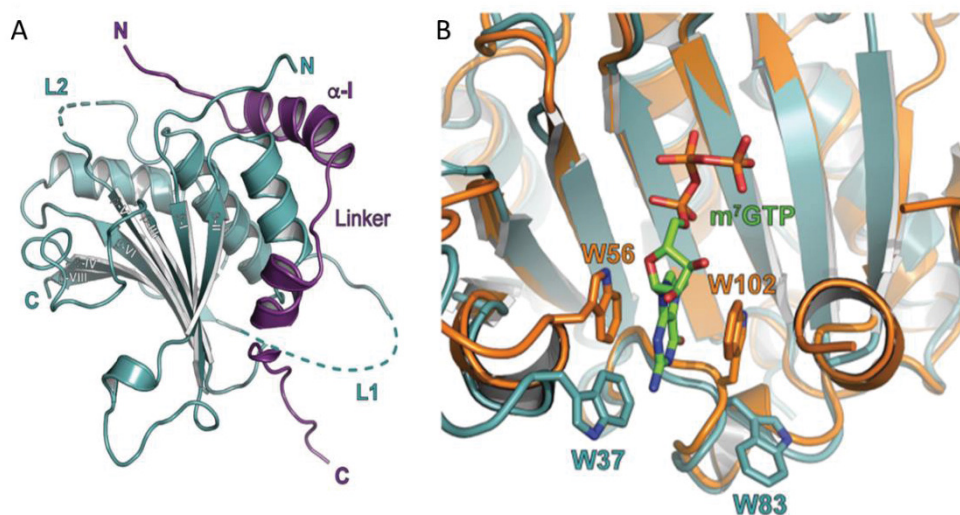


Figura 6 - Estrutura tridimensional de LmEIF4E1 (PDB: 5WB5). A) Em azul a estrutura de LmEIF4E1 em complexo com 4E-IP1 (em roxo). B) Sobreposição estrutura de LmEIF4E1 em azul e eIF4E1 humano em complexo com m⁷GTP (PDB: 5BXV) em laranja. A sobreposição destaca os resíduos do *stacking* de interação com m⁷GTP, mostrando a mobilidade dos resíduos W37 e W83 de LmEIF4E1 quando não há a interação com o análogo do cap. Fonte: MELEPPATTU, 2018.

A estrutura de LmEIF4E1 apresenta de forma geral a mesma configuração estrutural da família eIF4E. Na estrutura determinada os resíduos de triptofano correspondentes aos envolvidos na interação com o cap apresentam-se numa conformação “aberta” mostrando que tais resíduos possuem mobilidade entre o estado ligado e não ligado aos análogos do cap.

A estrutura de LmEIF4E1 foi determinada em complexo com o peptídeo de 4E-IP1. 4E-IP1 é uma proteína de 83 kDa que possui ortólogos entre os tripanossomatídeos, a interação entre LmEIF4E1 e 4E-IP1 é dada pela mesma região de interação (YXXXXLΦ em EIF4G e 4E-BP, e VXXFW em EIF4E) entre o fator 4E e o fator 4G. A formação do complexo LmEIF4E1-4E-IP1 reprime o início da tradução, bem como reduz a afinidade de LmEIF4E1 ao cap, indicando assim o seu papel de repressora da tradução, por competição da interação com o fator 4G (MELEPPATTU, 2018).

2 JUSTIFICATIVA

Apesar do grande número de estudos publicados envolvendo os fatores de iniciação da tradução, ainda existe muito a se determinar sobre o papel dos diferentes homólogos destes fatores na regulação da maquinaria de tradução. Em particular, o grande número de homólogos dos fatores eIF4E e eIF4G presentes em tripanossomatídeos e as interações específicas nas quais estão envolvidos sugerem vias de regulação diferenciadas presentes nesses organismos. Além disso, a estrutura única do cap do mRNA de tripanossomatídeos indica mecanismos de interação específicos com os fatores EIF4E, cujos detalhes atômicos ainda não foram elucidados.

Estes mecanismos de interações específicos do fator EIF4E com a estrutura única do cap-4 podem ser futuramente explorados para o desenvolvimento de drogas inibitórias para a terapia da Doença de Chagas, cujo tratamento atualmente é ineficiente e gera graves efeitos colaterais. Em vista disso, propomos nesse trabalho o estudo estrutural dos homólogos de eIF4E de *Trypanosoma cruzi* e suas interações com o cap e seus análogos.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Caracterização estrutural dos fatores de iniciação da tradução EIF4E de *Trypanosoma cruzi* e análise de suas interações com o cap do mRNA.

3.2 OBJETIVOS ESPECÍFICOS

- a) Produção recombinante de diferentes construções dos homólogos de EIF4E de *T. cruzi*;
- b) Avaliar a interação dos fatores EIF4E de *T. cruzi* com análogos do cap m⁷GTP e cap-4 utilizando a técnica de termoforese em microescala;
- c) Cristalizar os fatores EIF4E de *T. cruzi* e seus complexos com análogos do cap.
- d) Determinar a estrutura cristalográfica de EIF4E5 de *T. cruzi* em complexo com análogos do cap e analisar detalhadamente as interações.

4 MATERIAIS E MÉTODOS

4.1 DESENHO DAS CONSTRUÇÕES PARA EXPRESSÃO DOS FATORES

EIF4E DE *T. cruzi*

O desenho das construções baseou-se em análises estruturais anteriormente realizadas em nosso laboratório, a partir de modelos por homologia dos seis fatores EIF4E de *T. cruzi* (MATOS, 2018). Um esquema dos seis fatores EIF4E é mostrado na Figura 7. Regiões N e C-terminais preditas como não estruturadas foram truncadas, assim como deleções em regiões localizadas na vizinhança do sítio de interação com o cap, a fim de se analisar o papel das mesmas na interação com análogos do cap do mRNA.

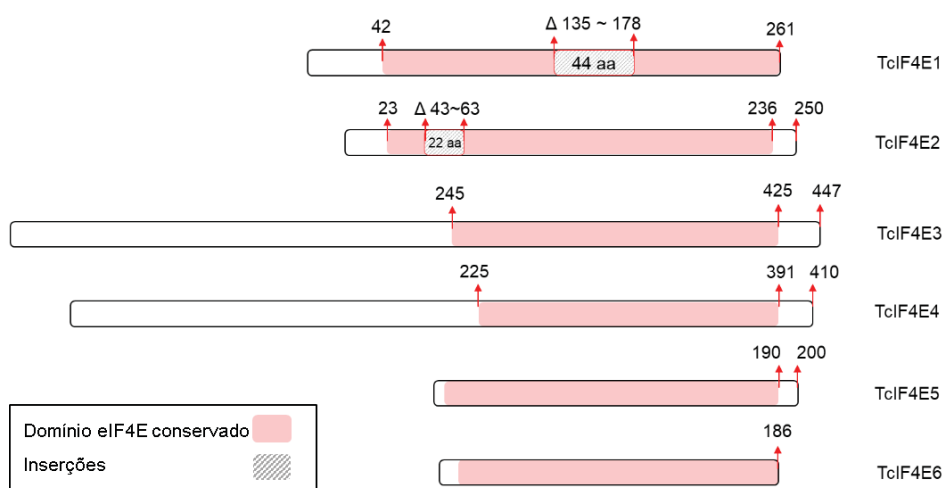


Figura 7 - Representação esquemáticas dos seis fatores EIF4E de *T. cruzi*.

As sequencias codificadoras para as diferentes construções foram adquiridas comercialmente (Biomatik), otimizadas para a expressão em *Escherichia coli* e clonadas no vetor pET28a (Novagen). Uma das construções do fator TcEIF4E1 foi cedida por colaboradores do projeto, clonada em vetor de expressão pDEST17. As construções possuem uma sequência de hexahistidina na região N-terminal, o que permite a realização de uma etapa de purificação por cromatografia de afinidade à níquel. Na tabela 3 estão descritas as 11 construções e algumas de suas características.

Identificação das construções	Descrição	Resíduos presentes	Peso molecular (kDa)
TcEIF4E1	Clonada em vetor pDEST17	1 – 261	31,6
TcEIF4E1ΔN	Deleção da extremidade N-terminal	42 - 261	27,2
TcEIF4E1ΔS4-H2	Deleção N-terminal e deleção dos resíduos 135 – 178 (loop que conecta a fita β4 e a hélice α2)	42 - 261; Δ 135 - 178	22,7
TcEIF4E2ΔN	N-terminal truncado	23 - 250	28,4
TcEIF4E2ΔS1-S2	N e C terminal truncados e deleção dos resíduos 42 – 63 (loop que conecta as fitas β1 e β2)	23 – 236; Δ 42 - 63	25,0
TcEIF4E3ΔN	N terminal truncado	244 - 447	25,2
TcEIF4E4ΔN	N terminal truncado	226 - 410	23,5
TcEIF4E5	Proteína inteira	1 – 200	24,8
TcEIF4E5ΔC	C terminal truncado	1 -189	23,7
TcEIF4E6	Proteína inteira	1 - 186	23,6

Tabela 3 - Descrição das construções dos fatores EIF4E de *T. cruzi* utilizados neste trabalho.

4.2 TESTES DE EXPRESSÃO

Os vetores de expressão derivados de pET28a, contendo as sequências codificadoras correspondentes às construções descritas na Tabela 3, foram transformados em *E. coli* por meio de choque térmico, as construções TcEIF4E2ΔN, TcEIF4E5 e TcEIF4E5ΔC em bactérias BL21(DE3) STAR e as construções, TcEIF4E1ΔN, TcEIF4E1ΔS4-H2, TcEIF4E2ΔNC, TcEIF4E2ΔS1-S2, TcEIF4E3ΔN, TcEIF4E4ΔN e TcEIF4E6 em células BL21 ΔSlyD pRARE2. Apesar da otimização dos genes para expressão em *E. coli*, as células BL21 ΔSlyD pRARE2 foram utilizadas devido à deleção da proteína SlyD, uma proteína rica em histidinas que aparece como um dos contaminantes principais após a cromatografia de afinidade à níquel.

As células foram plaqueadas em meio sólido LB (Lysogeny broth) contendo 25 µg/mL de Canamicina para as construções transformadas em BL21(DE3) STAR e em 25 µg/mL de Canamicina + 34 µg/mL de Cloranfenicol para as construções transformadas em BL21 ΔSlyD pRARE2. Colônias isoladas foram cultivadas em meio LB líquido (70 mL), contendo a mesma proporção de antibiótico da placa sólida e mantidos a 37°C, 200 RPM (rotações por minuto)

até atingirem a densidade ótica a 600 nm (D.O.600) entre 0,6 e 0,8. Do volume inicial da cultura, 50 mL foram aclimatados a 18°C e induzidos com 0,25 mM de IPTG e mantidos sob agitação por 16 horas para temperatura 18°C, os outros 20 mL da cultura inicial serviram como controle não induzido.

As culturas foram centrifugadas a 6000 x g por 15 minutos afim de retirar o meio LB, o pellet foi ressuspendido em tampão de lise (50 mM de Tris-HCl pH 8,0 (pH 8,8 para construção TcEIF4E2ΔN); 300 mM de NaCl; 20 mM de Imidazol). A lise das células foi por lisozima (150 µg/mL) e sonicação (amplitude 30%, com 10 ciclos de 10 segundos) em banho de gelo. O lisado foi então centrifugado a 20000 x g, a 4°C por 30 minutos afim de separar os extratos solúveis e insolúveis, a fração insolúvel foi ressuspendida em tampão de lise para a aplicação no gel de poliacrilamida. As amostras de TcEIF4E2ΔN, TcEIF4E5 e TcEIF4E5ΔC foram analisadas diretamente por eletroforese em gel de poliacrilamida (SDS-PAGE). As construções de TcEIF4E1ΔN, TcEIF4E1ΔS4-H2, TcEIF4E2ΔS1-S2, TcEIF4E3ΔN, TcEIF4E4ΔN e TcEIF4E6 foram analisadas após purificação de afinidade com resina suspensa (o protocolo está no anexo B), o qual foi realizado pela tecnologista da Plataforma de Purificação de Proteínas Tatiana Assunção.

4.3 EXPRESSÃO DAS CONSTRUÇÕES DE EIF4E DE *T. cruzi*

O cultivo para a expressão das proteínas foi realizado em meio líquido LB (de 500 mL a 1 litro, com 2% de pré-inóculo) a 37°C, 200 RPM, por um período de 2 a 3 horas até a D.O.600 entre 0,6 e 0,8. Após o crescimento as culturas foram aclimatadas por 30 minutos a 18°C, previamente à adição de 0,25 mM de IPTG. As culturas foram mantidas por aproximadamente 16 horas a 18°C sob agitação e então submetidas à centrifugação a 6000 x g por 15 minutos. O pellet de células foi ressuspendido em tampão A da cromatografia de afinidade, cada 500 mL de cultura foi ressuspendido em 10 mL do tampão (a composição dos tampões A está na Tabela 4).

Construção	Tampão A (afinidade)
TcEIF4E1, TcEIF4E5 e TcEIF4E5ΔC	50 mM de Tris-HCl pH 8; 0,2 M de NaCl; 0,2 M de Imidazol e 10 mM de β-mercaptoetanol
TcEIF4E2ΔN	50 mM de Tris-HCl pH 8,8; 0,2 M de NaCl; 0,2 M de Imidazol e 10 mM de βME
TcEIF4E1ΔN, TcEIF4E1ΔS4-H2 e TcEIF4E2ΔS1-S2	50 mM de Tris-HCl pH 8; 0,3 M de NaCl; 0,2 M de Imidazol e 10 mM de βME

Tabela 4 - Composição dos Tampões A de cromatografia de afinidade.

As amostras ressuspendidas foram incubadas com 150 µg/mL de lisozima, mantidas em gelo por aproximadamente 30 minutos, posteriormente foram lisadas em um microfluidificador (80 psi por 10 ciclos) (Microfluidics™). Após a lise foram separados os extratos solúveis e insolúveis por centrifugação a 20000 x g por 30 minutos.

4.4 PURIFICAÇÃO DAS PROTEÍNAS DE INTERESSE

4.4.1 Purificação em sistema automatizado

4.4.1.1 Cromatografia de afinidade à níquel

Todas as construções de TcEIF4E1, TcEIF4E2 e TcEIF4E5 foram purificadas utilizando um equipamento AKTA FPLC (GE Healthcare). A primeira etapa foi a cromatografia de afinidade a níquel, a amostra foi aplicada em uma coluna HisTrap HP 1 mL (GE Healthcare) equilibrada em tampão A e eluída em um gradiente de tampão B (composição dos tampões B está na Tabela 5)

Construção	Tampão B (afinidade)
TcEIF4E1, TcEIF4E5 e TcEIF4E5ΔC	50 mM de Tris-HCl pH 8; 0,2 M de NaCl; 0,5 M de Imidazol e 10 mM de βME
TcEIF4E2ΔN	50 mM de Tris-HCl pH 8,8; 0,2 M de NaCl; 0,5 M de Imidazol e 10 mM de βME
TcEIF4E1ΔN, TcEIF4E1ΔS4-H2 e TcEIF4E2ΔS1-S2	50 mM de Tris-HCl pH 8; 0,3 M de NaCl; 0,5 M de Imidazol e 10 mM de βME

Tabela 5 - Composição dos Tampões B utilizadas nas cromatografias de afinidade.

O método inicia com a aplicação da mostra na coluna, seguida de uma lavagem com 10 volumes de coluna (VC) com tampão A, o primeiro gradiente é

de 0 – 10% de tampão B de afinidade em 10 VC (gradiente de 0 – 15% para a construção TcEIF4E1) seguida de um segundo gradiente de tampão B de afinidade de 10 – 100% em 15 VC. As frações eluídas foram coletadas e as amostras foram analisadas pelo cromatograma gerado pelo equipamento e por SDS-PAGE.

4.4.1.2 Cromatografia por troca iônica

As construções TcEIF4E1ΔN e TcEIF4E1ΔS4-H2 foram as únicas a serem submetidas a uma cromatografia de troca iônica. Após a análise das frações eluídas na cromatografia de afinidade, as frações mais puras foram selecionadas e reunidas, então dialisadas em tampão A de troca iônica (composição do tampão A: 50 mM de Tris HCl pH8, 50 mM de NaCl e 2 mM de DTT). Após a dialise a amostra foi centrifugada a 15000 x g por 10 minutos para a retirada de precipitados. Utilizou-se a coluna HiTrap QHP 5 mL (GE Healthcare). Após a injeção da amostra a coluna é lavada com 3 VC de tampão A, seguida de um gradiente de eluição linear em 20 VC de tampão B (50 mM Tris HCl pH 8; 1 M de NaCl e 2 mM de DTT). As frações eluídas foram coletadas e analisadas pelo cromatograma gerado pelo equipamento e por SDS-PAGE.

4.4.1.3 Cromatografia de exclusão por tamanho

As construções TcEIF4E1, TcEIF4E2ΔN, TcEIF4E2ΔS1-S2, TcEIF4E5 e TcEIF4E5ΔC após a purificação de afinidade foram purificadas por cromatografia de exclusão por tamanho e as construções TcEIF4E1ΔN e TcEIF4E1ΔS4-H2 passaram após a cromatografia de afinidade pela cromatografia de troca iônica e posteriormente seguiram para a cromatografia de exclusão por tamanho.

As construções TcEIF4E1, TcEIF4E1ΔN, TcEIF4E1ΔS4-H2, TcEIF4E2ΔN, TcEIF4E2ΔS1-S2, TcEIF4E5 e TcEIF4E5ΔC tiveram suas amostras selecionadas, reunidas e concentradas (concentrador Amicon 3000 MWCO, 4000 x g a 4°C) a um volume de 2,5 mL e centrifugada a 15000 x g por 10 minutos a 4°C, antes da aplicação em coluna Superdex 75 16/600 (GE Healthcare). A cromatografia ocorre de forma isocrática em 1 VC, com tampão

contendo: 20 mM de Tris-HCl pH 8 (pH8,8 para TcEIF4E2ΔN); 150 mM de NaCl e 2 mM de DTT. As frações eluídas foram coletadas e analisadas pelo cromatograma gerado pelo equipamento e por SDS-PAGE.

As construções TcEIF4E3ΔN, TcEIF4E4ΔN e TcEIF4E6 passaram pela cromatografia de exclusão por tamanho após a purificação por gravidade, após a seleção das amostras, as mesmas foram reunidas e concentradas a um volume de 500 μL e centrifugada a 15000 x g por 10 minutos a 4°C, antes da aplicação em coluna Superdex 75 Increase 10/300 (GE Healthcare). A cromatografia ocorre de forma isocrática em 1 VC, com tampão contendo: 50 mM de Tris-HCl pH 8; 300 mM de NaCl e 2 mM de DTT. As frações eluídas foram coletadas e analisadas pelo cromatograma gerado pelo equipamento e por SDS-PAGE.

4.4.2 Purificação por gravidade

As construções de TcEIF4E3ΔN, TcEIF4E4ΔN e TcEIF4E6 foram purificadas por gravidade em resina suspensa. A amostra foi incubada com 100 μL de resina (NiSepharose 6 Fast Flow da GE Healthcare) por 4 horas a 4°C sob agitação leve. Os tampões utilizados são descritos na Tabela 6.

Tampão A	50 mM Tris pH 8; 0,3 M NaCl; 20 mM imidazol ; 10 mM βME
Tampão de lavagem 1	50 mM Tris pH 8; 0,3 M NaCl; 50 mM imidazol ; 10 mM βME
Tampão de lavagem 2	50 mM Tris pH 8; 0,3 M NaCl; 60 mM imidazol ; 10 mM βME
Tampão de lavagem 3	50 mM Tris pH 8; 0,3 M NaCl; 70 mM imidazol ; 10 mM βME
Tampão de lavagem 4	50 mM Tris pH 8; 0,3 M NaCl; 80 mM imidazol ; 10 mM βME
Tampão de eluição B1	50 mM Tris pH 8; 0,3 M NaCl; 100 mM imidazol ; 10 mM βME
Tampão de eluição B2	50 mM Tris pH 8; 0,3 M NaCl; 200 mM imidazol ; 10 mM βME
Tampão de eluição B3	50 mM Tris pH 8; 0,3 M NaCl; 500 mM imidazol ; 10 mM βME

Tabela 6 - Tampões utilizados na purificação por gravidade.

O método usado segue:

- 3 lavagens com 1 mL de tampão A (30 VC no total)
- 2 lavagens com 1 mL de tampão de lavagem 1 (20 VC no total)
- 1 lavagem com 1 mL tampão de lavagem 2 (10 VC no total)

- 1 lavagem com 1 mL tampão de lavagem 3 (10 VC no total)
- 1 lavagem com 1 mL tampão de lavagem 4 (10 VC no total)
- 2 eluições com 300 μ L de tampão B1 (3 VC cada eluição)
- 2 eluições com 300 μ L de tampão B2 (3 VC cada eluição)
- 2 eluições com 300 μ L de tampão B3 (3 VC cada eluição)

As frações de lavagem e as frações eluídas foram coletadas e analisadas por SDS-PAGE.

4.5 ENSAIO DE INTERAÇÃO POR TERMOFORESE EM MICRO-ESCALA

O equipamento utilizado para o ensaio de termoforese em micro escala foi o Monolith NT.115 (NanoTemper). Foram testados 2 análogos do cap: m⁷GTP (Jena Bioscience) e cap-4. O cap-4 foi sintetizado por nossas colaboradoras Dra. Sylvie Pochet e Dra. Sophie Vichier-Guerre do Instituto Pasteur – Paris, segundo método descrito em Reolon et al. 2019 (Anexo C). A metodologia aplicada requer a marcação fluorescente de uma das moléculas parceiras na interação enquanto a outra é titulada. Assim, as proteínas foram marcadas utilizando-se o kit RED-tris-NTA (NanoTemper) para marcação não covalente da cauda de hexahistidina. O tampão utilizado no ensaio foi 20 mM de Tris HCl pH8 (pH8,8 para TcEIF4E2 Δ N); 100 mM de NaCl, 2 mM de DTT e 0,05% de Tween 20.

Inicialmente foi realizado um pré-teste de fluorescência com as proteínas, para avaliação da marcação e casos de agregação da mesma com os capilares utilizados no ensaio. Os ensaios de interação foram realizados em triplicata, os ligantes sendo titulados a partir de uma concentração máxima de 2,5 mM para os ensaios com TcEIF4E2 Δ N, TcEIF4E2 Δ S1-S2, TcEIF4E3 Δ N, TcEIF4E5, TcEIF4E5 Δ C e TcEIF4E6; e concentração máxima de 0,5 mM para os ensaios com TcEIF4E1, TcEIF4E1 Δ N e TcEIF4E1 Δ S4-H2. A proteína marcada foi utilizada em concentração final constante de 12,5 nM.

Os dados foram analisados utilizando-se o software MO. Affinity Analysis v.2.2.7 (NanoTemper). O software normaliza os dados obtidos no ensaio, relação entre a fluorescência durante a exposição ao gradiente de temperatura e a fluorescência inicial anterior ao gradiente. A partir da curva de variação de

fluorescência normalizada em função da concentração do ligante é calculada a constante de dissociação assim como os erros associados e parâmetros como sinal sobre ruído.

4.6 ENSAIO DE CRISTALIZAÇÃO

Após purificação TcEIF4E1ΔN, TcEIF4E1ΔS4-H2, TcEIF4E2ΔN e TcEIF4E5 foram submetidas ao ensaio de cristalização, utilizando o método de difusão a vapor, com os kits comerciais JCSG+ (Molecular Dimensions), PACT *premier* (Molecular Dimensions) e Index^{HT} (Hampton Research), para a produção das gotas foi utilizado o equipamento Mosquito (TPP LabTech).

4.6.1 TcEIF4E2ΔN

A construção TcEIF4E2ΔN foi utilizada em concentração próxima de 1,5 mg/mL em tampão 20 mM Tris HCl pH 8,8; 150 mM NaCl e 2 mM de DTT. O primeiro teste foi feito com os kits comerciais JCSG+ e PACT *premier*, com a proporção de proteína/reservatório de 1:2; 1:1 e 2:1 com volume final da gota de 600 nL.

Um segundo teste foi realizado após a incubação com os análogos cap-4 e m⁷GTP a proteína. Foi adicionado um excesso molar de 2 vezes a concentração da proteína (incubação por 30 min), TcEIF4E2ΔN em concentração próxima a 0,9 mg/mL foi testada com os kits JCSG+ e Index^{HT}, com a proporção de 1:1 de proteína/reservatório em um volume final de 400 nL.

4.6.2 TcEIF4E1ΔN e TcEIF4E1ΔS4-H2

As construções TcEIF4E1ΔN e TcEIF4E1ΔS4-H2 foram inicialmente testadas com os kits comerciais JCSG+ e Index^{HT}, ambas em tampão 20 mM Tris HCl pH 8; 150 mM NaCl e 1 mM de DTT e a uma concentração em torno de 5 mg/mL. Ambas proteínas foram incubadas com os análogos do cap (cap-4 e m⁷GTP) em um excesso molar de 2 vezes a concentração da proteína. Da construção TcEIF4E1ΔS4-H2 foram feitas gotas de proporção 1:1 de proteína/reservatório com um volume final da gota de 340 nL. E da construção

TcEIF4E1ΔN a mesma proporção 1:1 foi mantida, mas o volume final da gota era de 360 nL.

Algumas condições do teste com TcEIF4E1ΔN que apresentaram precipitados cristalinos foram refinadas. No primeiro refinamento as condições variaram conforme a Tabela 7, com a concentração da proteína de 5mg/mL.

← Concentração de PEG 3350	pH →			
	0,1 M Bis-Tris pH 5,5 19% PEG 3350	0,1 M BIS-TRIS pH 6 19% PEG 3350	0,1 M BIS-TRIS pH 6,5 19% PEG 3350	0,1 M HEPES pH 7 19% PEG 3350
	0,1 M Bis-Tris pH 5,5 22% PEG 3350	0,1 M BIS-TRIS pH 6 22% PEG 3350	0,1 M BIS-TRIS pH 6,5 22% PEG 3350	0,1 M HEPES pH 7 22% PEG 3350
	0,1 M Bis-Tris pH 5,5 25% PEG 3350	0,1 M BIS-TRIS pH 6 25% PEG 3350	0,1 M BIS-TRIS pH 6,5 25% PEG 3350	0,1 M HEPES pH 7 25% PEG 3350
	0,1 M Bis-Tris pH 5,5 28% PEG 3350	0,1 M BIS-TRIS pH 6 28% PEG 3350	0,1 M BIS-TRIS pH 6,5 28% PEG 3350	0,1 M HEPES pH 7 28% PEG 3350

Tabela 7 - Condições de cristalização utilizados no primeiro refinamento de TcEIF4E1ΔN.

No segundo refinamento foi testado duas concentrações da proteína, para a concentração de 4,8 mg/mL as condições variaram conforme a Tabela 8.

→ pH	→ concentração de PEG 3350				→ concentração de PEG 3350			
	0,1M Bis-Tris pH6 22% PEG 3350	0,1M Bis-Tris pH6 25% PEG 3350	0,1M Bis-Tris pH6 28% PEG 3350	0,1M Bis-Tris pH6 30% PEG 3350	0,1M Bis-Tris pH6 22% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6 25% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6 28% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6 30% PEG 3350 0,2M Ac. de Amônio
	0,1M Bis-Tris pH6,5 22% PEG 3350	0,1M Bis-Tris pH6,5 25% PEG 3350	0,1M Bis-Tris pH6,5 28% PEG 3350	0,1M Bis-Tris pH6,5 30% PEG 3350	0,1M Bis-Tris pH6,5 22% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6,5 25% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6,5 28% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6,5 30% PEG 3350 0,2M Ac. de Amônio

Tabela 8 - Condições de cristalização utilizados no segundo refinamento de TcEIF4E1ΔN, com a concentração de 4,8 mg/mL.

E para a concentração de 4 mg/mL as condições variaram conforme a Tabela 9.

		→ concentração de PEG 3350				→ concentração de PEG 3350			
pH ↓	0,1M Bis-Tris pH6 25% PEG 3350	0,1M Bis-Tris pH6 28% PEG 3350	0,1M Bis-Tris pH6 30% PEG 3350e	0,1M Bis-Tris pH6 32% PEG 3350	0,1M Bis-Tris pH6 25% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6 28% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6 30% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6 32% PEG 3350 0,2M Ac. de Amônio	
	0,1M Bis-Tris pH6,5 25% PEG 3350	0,1M Bis-Tris pH6,5 28% PEG 3350	0,1M Bis-Tris pH6,5 30% PEG 3350	0,1M Bis-Tris pH6,5 32% PEG 3350	0,1M Bis-Tris pH6,5 25% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6,5 28% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6,5 30% PEG 3350 0,2M Ac. de Amônio	0,1M Bis-Tris pH6,5 32% PEG 3350 0,2M Ac. de Amônio	

Tabela 9 – Condições de cristalização utilizados no segundo refinamento de TcEIF4E1ΔN, com a concentração de 4 mg/mL.

4.6.3 TcEIF4E5

TcEIF4E5 foi utilizada a aproximadamente 5,4 mg/mL em tampão 20 mM Tris HCl pH 8, 150 mM NaCl e 2 mM de DTT, foram testados inicialmente os kits comerciais JCSG+, PACT *premier* e Index^{HT}. Observou-se cristais no kit Index^{HT} (condição H4: 0,2 M citrato de amônia pH 7 + 20% w/v PEG 3350) e no kit PACT *premier* (condição D11: 0,2 M cloreto de magnésio hexahidratado + 0,1 M tris HCl pH 8 + 20% w/v PEG 6000). Assim foram realizados refinamentos a fim de determinar as melhores condições de cristalização para a proteína.

Para o primeiro refinamento em ambas as condições foram utilizados 17%, 20% e 23% dos respectivos PEG, para a condição H4 foi utilizado acetato de amônia com a variação de pH entre 6,6 e 7,2 e para a condição D11 foram testados os pH 7; 7,5; 8; e 8,5.

Um segundo refinamento foi realizado com base nas mesmas condições anteriores, porém desta vez a variação dos respectivos PEG foi entre 15, 17, 20, 23 e 26%. Para a condição H4 foi utilizado o citrato de amônia variando entre pH 6,5; 7 e 7,5. E para a condição D11 a variação do pH foi entre 7,5; 8 e 8,5. Neste segundo refinamento foi utilizado a proteína em concentrações de 5,8; 7,7 e 11,8 mg/mL.

Um terceiro refinamento foi realizado, e nesta etapa a proteína foi incubada com os análogos do cap (cap-1 e cap-4), foi adicionado um excesso molar de 2 vezes a concentração da proteína (5,5 mg/mL). As condições

testadas variaram entre 0,1 M Bis-Tris pH 6 e 6,5 + PEG 3350 (28%, 30%, 32% e 34%) + 0,2 M de Sulfato de Lítio.

5 RESULTADOS E DISCUSSÃO

5.1 TESTES DE EXPRESSÃO

5.1.1 Expressão em cepa *E. coli* BL21(DE3) STAR

A expressão em cepa *E. coli* BL21(DE3) STAR em temperatura de 18°C por 16 horas resultou na produção da proteína majoritariamente na fração solúvel da construção TcEIF4E2ΔN (Figura 8) e para todas as construções de TcEIF4E5 (Figura 9). No caso das construções de TcEIF4E5 observa-se uma expressão basal da proteína recombinante também no extrato não induzido com IPTG, porém esta expressão é aumentada com a indução.

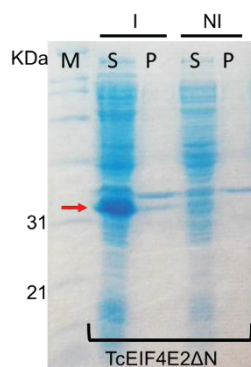


Figura 8 – Análise dos testes de expressão TcEIF4E2ΔN (28 kDa) em SDS-PAGE. M: marcador de massa molecular (*Bio-Rad Broad Range*); NI: controle não induzido; I: expressão induzida com IPTG; S: fração solúvel; P: (precipitado) fração insolúvel. A banda de expressão de TcEIF4E2ΔN apresenta migração diferenciada em gel de eletroforese, ligeiramente acima do seu peso molecular. A seta vermelha indica a proteína de interesse.

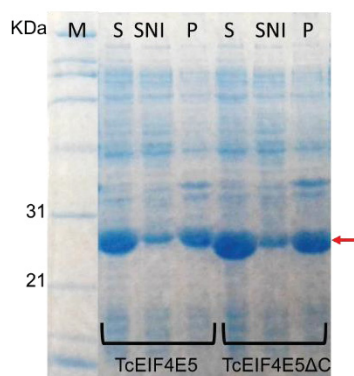


Figura 9 – Análise dos testes de expressão das construções TcEIF4E5 (24 kDa) e TcEIF4E5ΔC (23 kDa) em SDS-PAGE. M: marcador de massa molecular (*Bio-Rad Broad Range*); SNI: fração solúvel do controle não induzido; S: fração solúvel; P: (precipitado) fração insolúvel. A seta vermelha indica a proteína de interesse.

Para as demais construções cujo nível de expressão na fração solúvel não foi satisfatório, um segundo teste foi realizado na cepa de *E. coli* BL21 Δ SlyD pRARE2. O plasmídeo pRARE2 contém genes que codificam para a produção de tRNAs de códons raros, a cepa ainda possui deleção no gene da isômerase SlyD, um contaminante rico em histidinas, sendo um possível contaminante nas purificações por afinidade a níquel.

5.1.2 Expressão e teste de purificação em cepa *E. coli* BL21 Δ SlyD pRARE2

Os testes de expressão e purificação preliminares das construções TcEIF4E1 Δ N, TcEIF4E1 Δ S4-H2, TcEIF4E2 Δ NC, TcEIF4E2 Δ S1-S2, TcEIF4E3 Δ N, TcEIF4E4 Δ N e TcEIF4E6 foram realizadas pela colaboradora do projeto Tatiana Assunção (o protocolo utilizado está no Anexo A) o resultado se encontra nas Figuras 10, 11, 12 e 13.

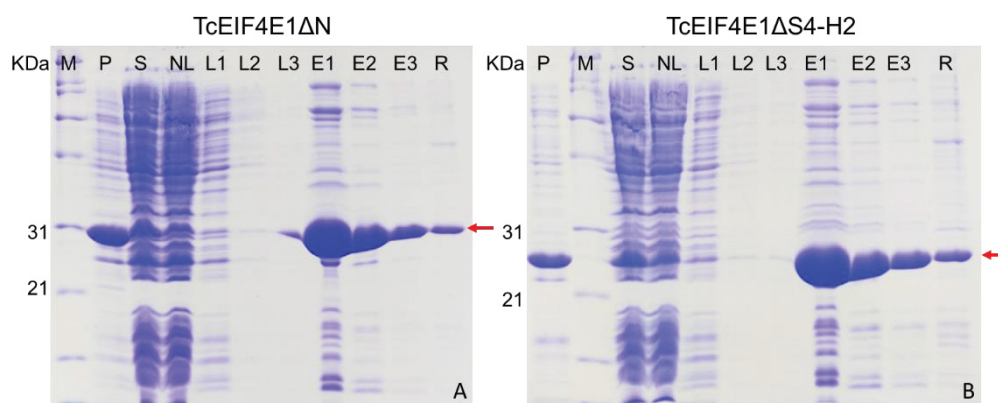


Figura 10 – Análise dos resultados dos testes de purificação das construções TcEIF4E1 Δ N e TcEIF4E1 Δ INS por SDS-PAGE. Pesos moleculares: TcEIF4E1 Δ N-27 KDa e TcEIF4E1 Δ INS-22 KDa. M: marcador de peso molecular (Bio-Rad Broad Range), P: (precipitado) fração insolúvel, S: fração solúvel, NL: fração não ligada na resina, L1, L2 e L3: lavagem 1, 2 e 3 respectivamente, E1, E2 e E3: eluição com tampão B 1, 2 e 3 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

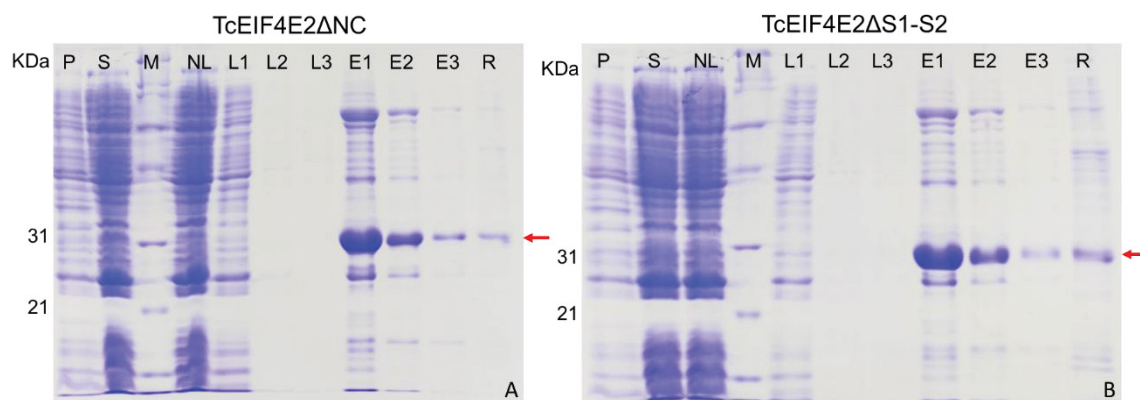


Figura 11 - Análise dos resultados dos testes de purificação das construções TcEIF4E2 Δ NC e TcEIF4E2 Δ S1-S2 por SDS-PAGE. Pesos moleculares: TcEIF4E2 Δ NC - 27 KDa e TcEIF4E2 Δ S1-S2 - 25 KDa. M: marcador de peso molecular (Bio-Rad Broad Range), P: (precipitado) fração insolúvel, S: fração solúvel, NL: fração não ligada na resina, L1, L2 e L3: lavagem 1, 2 e 3 respectivamente, E1, E2 e E3: eluição com tampão B 1, 2 e 3 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

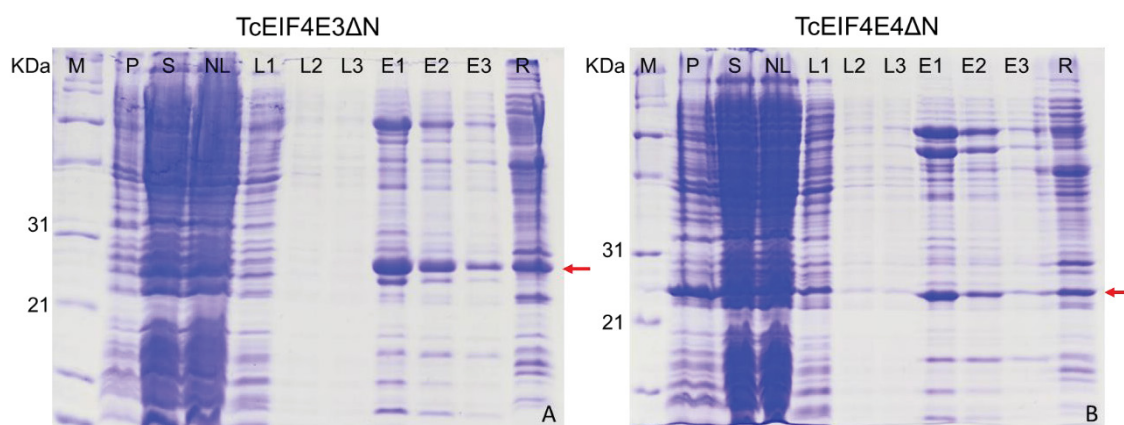


Figura 12 - Análise dos resultados dos testes de purificação das construções TcEIF4E3 Δ N e TcEIF4E4 Δ N por SDS-PAGE. Pesos moleculares: TcEIF4E3 Δ N-25 KDa e TcEIF4E4 Δ N-23,5 KDa. M: marcador de peso molecular (Bio-Rad Broad Range), P: (precipitado) fração insolúvel, S: fração solúvel, NL: fração não ligada na resina, L1, L2 e L3: lavagem 1, 2 e 3 respectivamente, E1, E2 e E3: eluição com tampão B 1, 2 e 3 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

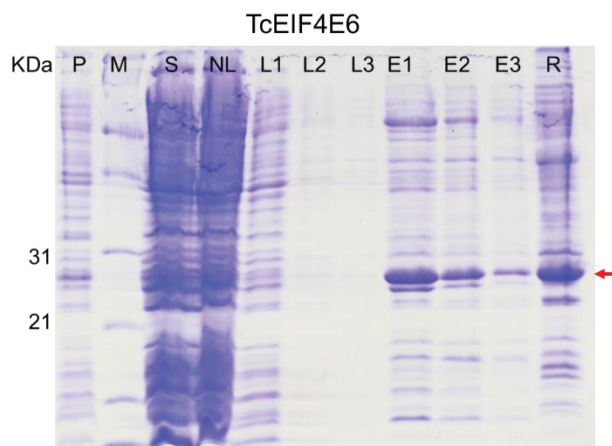


Figura 13 - Análise do resultado do teste de purificação da construção TcEIF4E6 por SDS-PAGE. Peso molecular: TcEIF4E6 - 23,5 KDa. M: marcador de peso molecular (Bio-Rad Broad Range), P: (precipitado) fração insolúvel, S: fração solúvel, NL: fração não ligada na resina, L1, L2 e L3: lavagem 1, 2 e 3 respectivamente, E1, E2 e E3: eluição com tampão B 1, 2 e 3 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

A partir destes resultados preliminares, que nos permitiu observar as proteínas de interesse nas frações de eluição da purificação por afinidade à níquel, estas foram submetidas à expressão e purificação em média escala. As construções TcEIF4E1 Δ N, TcEIF4E1 Δ S4-H2 e TcEIF4E2 Δ S1-S2 foram purificadas em sistema automatizado, e as construções TcEIF4E3 Δ N, TcEIF4E4 Δ N e TcEIF4E6 foram purificadas por gravidade.

5.2 EXPRESSÃO E PURIFICAÇÃO

5.2.1 TcEIF4E1, TcEIF4E1 Δ N e TcEIF4E1 Δ S4-H2

5.2.1.1 TcEIF4E1

A purificação da construção TcEIF4E1 partiu de 1 litro de cultura bacteriana e constituiu de dois passos, pela cromatografia de afinidade a níquel e pela cromatografia de exclusão por tamanho. Devido ao bom rendimento da purificação por afinidade a níquel não foi necessário a etapa de purificação por troca iônica. Os cromatogramas e a eletroforese em gel de poli(acrilamida) da purificação de afinidade são mostrados na Figura 14.

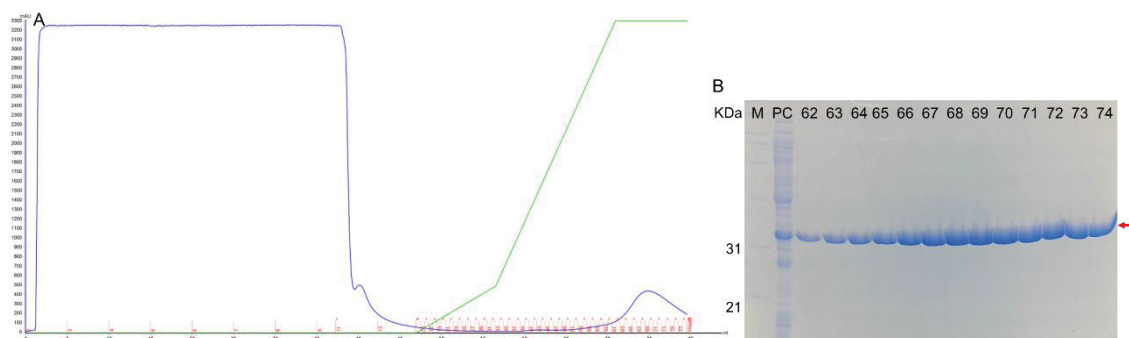


Figura 14 - Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1 por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; 62 – 74: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A partir da purificação por afinidade a níquel, as frações 64 - 76 foram reunidas e concentradas para a cromatografia de exclusão por tamanho.

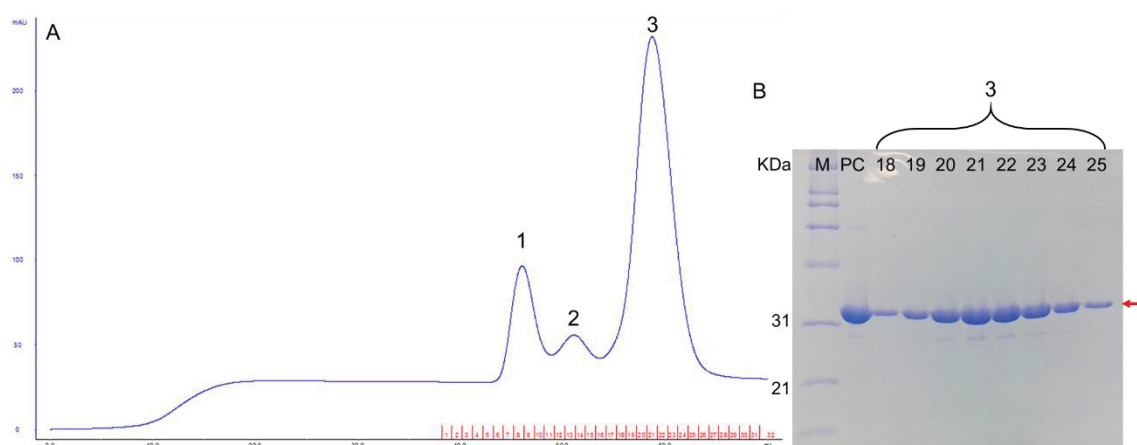


Figura 15 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1 por exclusão de tamanho em coluna Superdex 75 16/600. A) a linha azul corresponde à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; 18 – 25: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A cromatografia de exclusão por tamanho gerou 3 picos, os picos 1 e 2 correspondem a estados oligoméricos da própria proteína, conforme análise em SDS-PAGE (não mostrado). O volume de eluição correspondente ao pico 3, em torno de 60 mL condiz com o estado monomérico da proteína segundo o manual da coluna. O rendimento da purificação foi suficiente para realizar os ensaios de interação por termoforese em microescala, sendo utilizada a fração 23 da eluição.

5.2.1.2 TcEIF4E1ΔN

A construção TcEIF4E1ΔN foi purificada a partir de 1 litro de cultura bacteriana e constituiu de três passos, cromatografia de afinidade a níquel, cromatografia de troca iônica e cromatografia de exclusão por tamanho. O cromatograma e a eletroforese em gel de poliacrilamida da purificação de afinidade a níquel são mostrados na Figura 16.

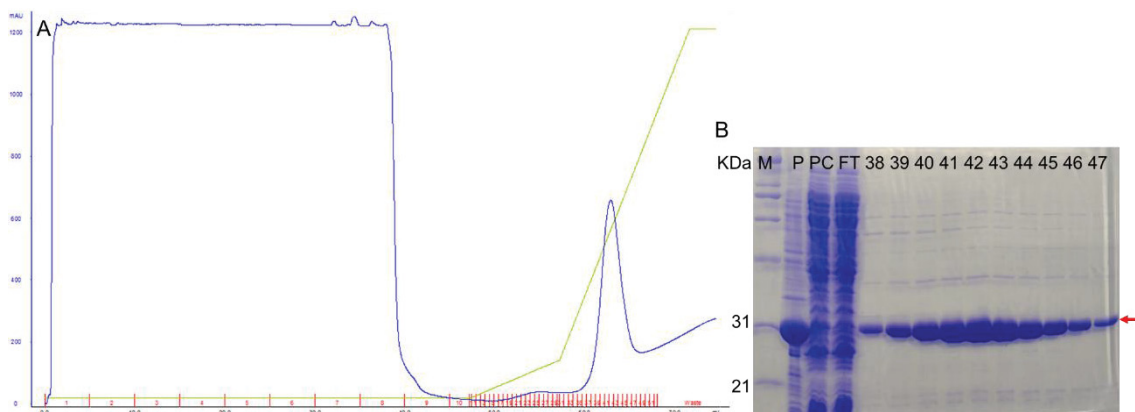


Figura 16 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔN por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); P: pellet (sedimentado); PC: amostra pré-coluna; FT: flowthrough; 38 – 47: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A partir da cromatografia de afinidade a níquel, as frações 38 - 48 foram reunidas e dialisadas para a cromatografia de troca aniônica. A cromatografia de troca iônica foi realizada a fim de retirar os contaminantes visíveis no SDS-PAGE.

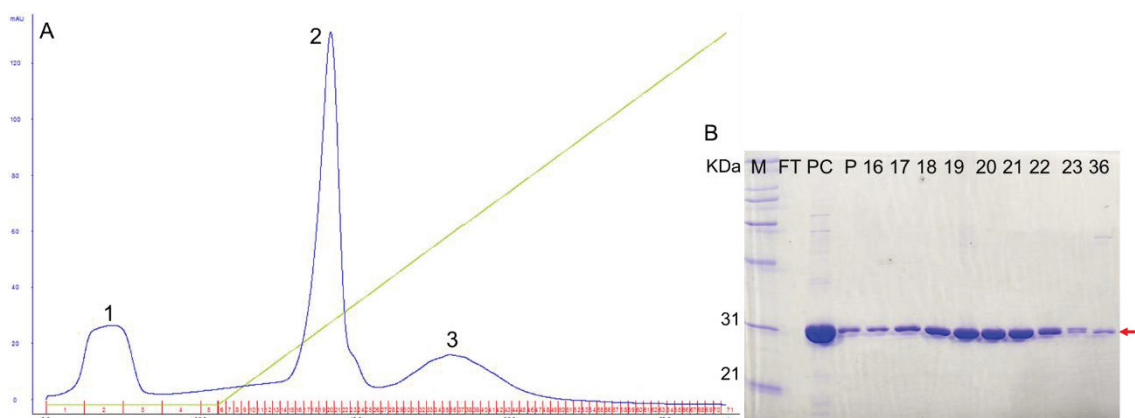


Figura 17 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔN por troca iônica. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); FT: flowthrough; PC: amostra pré-coluna. A seta vermelha indica a proteína de interesse.

O pico 2 corresponde as bandas da proteína no SDS-PAGE da Figura 16B, o pico 1 e o pico 3 também correspondem a proteína de interesse (dados não mostrados), correspondendo a populações diferentes da proteína. A partir da cromatografia de troca iônica foram concentradas as frações do pico majoritário 17 - 22 para a cromatografia de exclusão por tamanho (Figura 18).

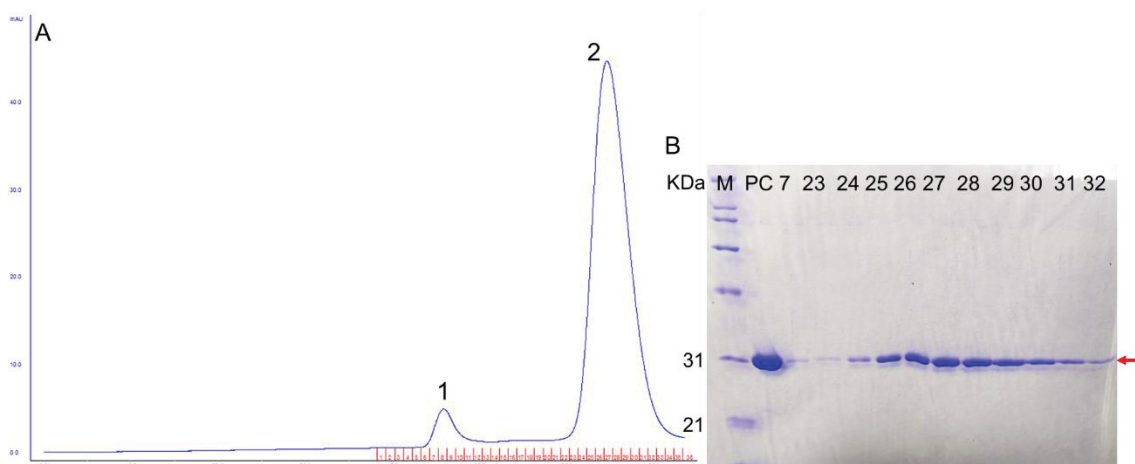


Figura 18 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔN por exclusão de tamanho em coluna Superdex 75 16/600. A) a linha azul corresponde à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna. A seta vermelha indica a proteína de interesse.

O pico 1 eluído em torno de 40 – 50 mL corresponde a um estado agregado da proteína, e o pico 2 eluído em torno de 60 – 70 mL corresponde ao estado monomérico da proteína. O rendimento foi suficiente para a realização dos ensaios de interação por termoforese em microescala, bem como para os ensaios de cristalização.

5.2.1.3 TcEIF4E1ΔS4-H2

A purificação de TcEIF4E1ΔS4-H2 constituiu de três etapas de purificação: afinidade a níquel, troca iônica e exclusão por tamanho. A purificação foi realizada com 1 litro de cultura bacteriana. A Figura 19 apresenta a cromatografia de afinidade a níquel e a eletroforese em gel de poliacrilamida.

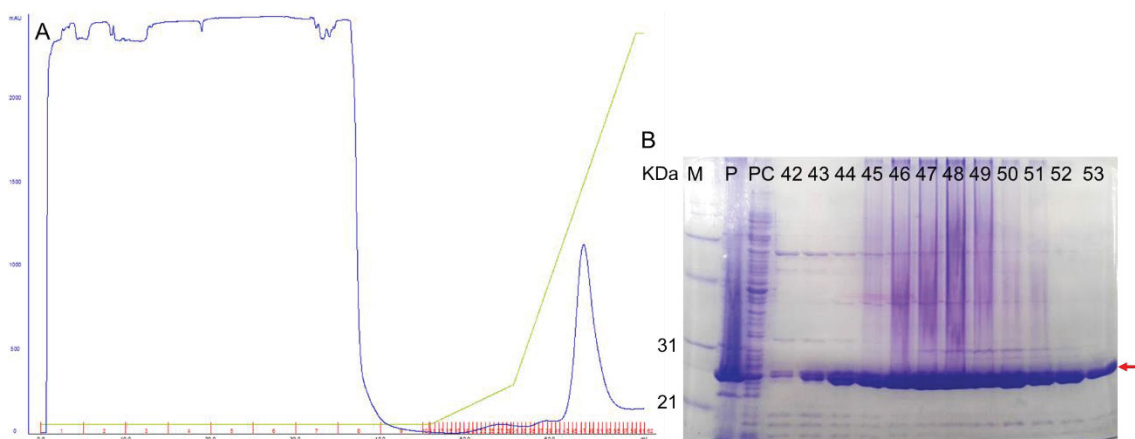


Figura 19 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔS4-H2 por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); P: pellet (sedimentado); PC: amostra pré-coluna; 42 – 53: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A partir da purificação de afinidade a níquel foram dialisadas as frações 43 – 53 para a cromatografia de troca iônica. A Figura 20 apresenta o cromatograma e a eletroforese em gel de poliacrilamida.

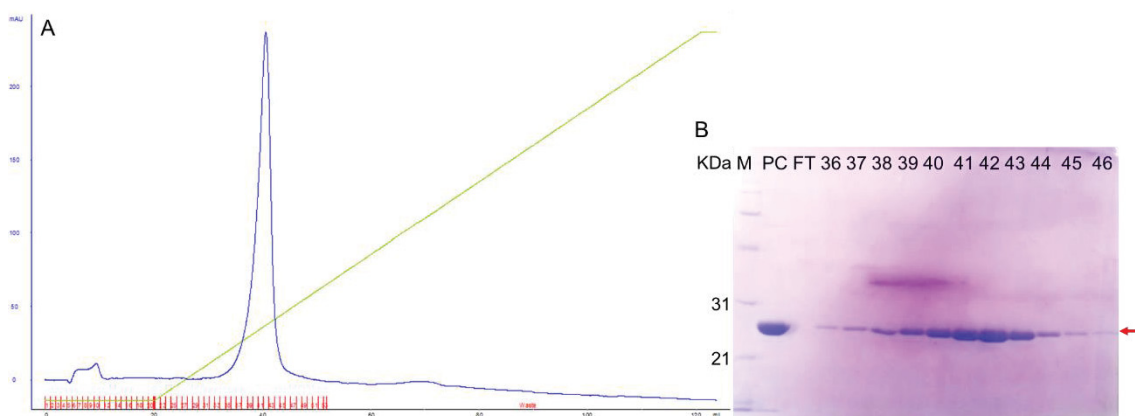


Figura 20 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1ΔS4-H2 por troca iônica. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; FT: flowthrough. A seta vermelha indica a proteína de interesse.

A partir da cromatografia de troca iônica as frações 38 – 45 foram concentradas para a purificação de exclusão por tamanho.

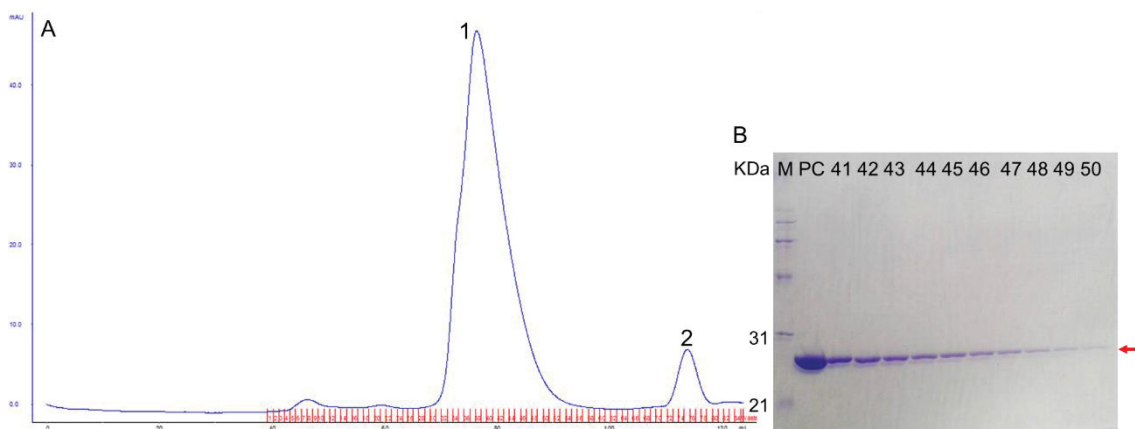


Figura 21 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E1 Δ S4-H2 por exclusão de tamanho, em coluna Superdex 75 16/600. A) a linha azul corresponde à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna. A seta vermelha indica a proteína de interesse.

O pico 1 corresponde ao estado monomérico da proteína, eluído em torno de 80 mL, o pico 2 corresponde a contaminantes de tamanho menor que o da proteína, e que foram separados na cromatografia de exclusão por tamanho. Para o ensaio de termoforese em microescala a fração 48 da gel filtração foi utilizada. E um ensaio de cristalização foi realizado com a concentração das frações 37 - 47, porém este ensaio não resultou em nenhum cristal e nenhuma condição promissora foi observada.

5.2.2 TcEIF4E2 Δ N e TcEIF4E2 Δ S1-S2

5.2.2.1 TcEIF4E2 Δ N

A construção TcEIF4E2 Δ N foi expressa em 500 mL de cultura bacteriana, e a purificação constituiu de duas etapas: cromatografia de afinidade a níquel e exclusão por tamanho, o resultado da cromatografia de afinidade a níquel é mostrado na Figura 22.

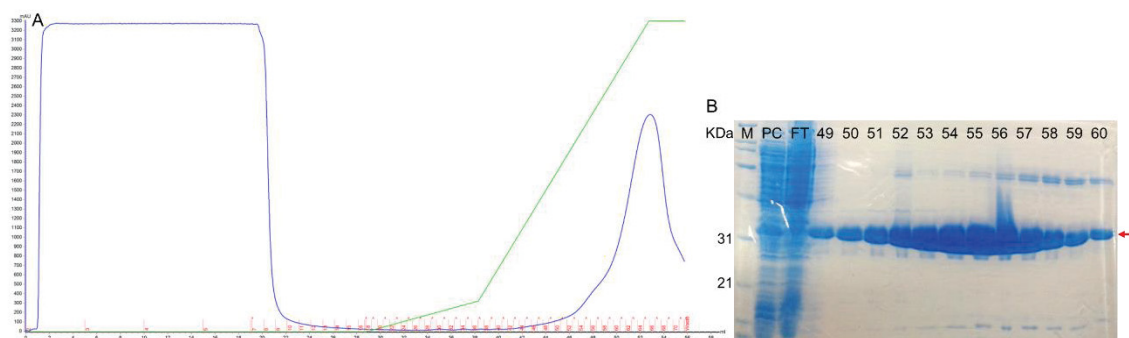


Figura 22 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E2ΔN por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; FT: flowthrough; 49 – 60: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

Observa-se a precipitação da proteína quando esta é armazenada em banho de gelo mesmo por curto período (a partir de uma hora), o que indica uma baixa solubilidade da mesma. Apesar da perda de uma fração da proteína por precipitação as frações 51 – 58 da cromatografia de afinidade foram concentradas e submetidas à cromatografia de exclusão por tamanho.

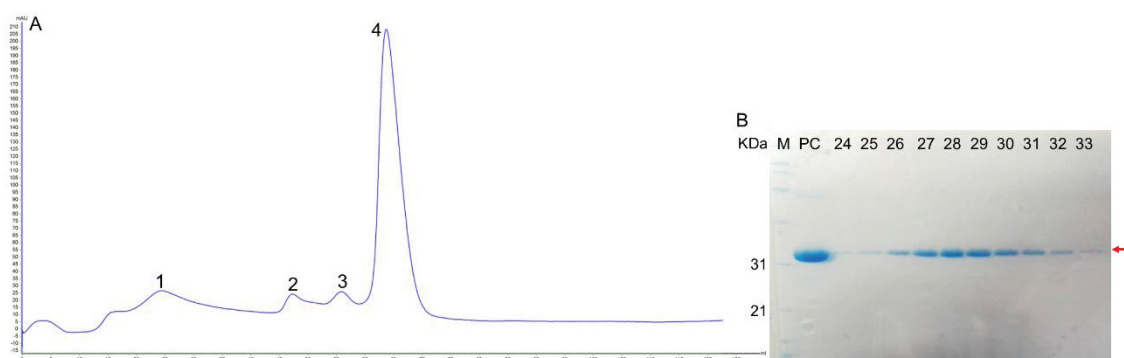


Figura 23 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E2ΔN por exclusão de tamanho em coluna Superdex 75 16/600. A) a linha azul corresponde à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna. A seta vermelha indica a proteína de interesse.

O pico majoritário da cromatografia de exclusão por tamanho, eluído em 65 mL corresponde ao estado monomérico da proteína e os picos 1, 2 e 3 correspondem aos contaminantes de maior peso molecular que ao da proteína de interesse. As frações 27 – 30 foram concentradas e utilizadas no ensaio de cristalização.

5.2.2.2 TcEIF4E2 Δ S1-S2

A construção TcEIF4E2 Δ S1-S2 foi expressa em 1 L de cultura bacteriana, e a purificação constituiu de duas etapas: cromatografia de afinidade a níquel e exclusão por tamanho, o resultado da cromatografia de afinidade a níquel é mostrado na Figura 24.

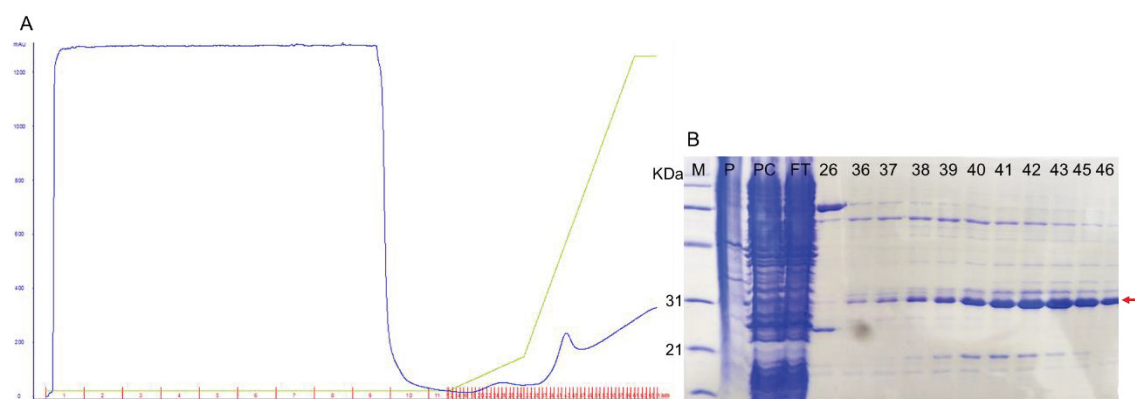


Figura 24 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E2 Δ S1-S2 por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorvância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); P: pellet (sedimentado); PC: amostra pré-coluna; FT: flowthrough; 26, 36 – 46: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A partir da purificação por afinidade a níquel as frações 38 – 47 foram reunidas e concentradas para a cromatografia de exclusão por tamanho.

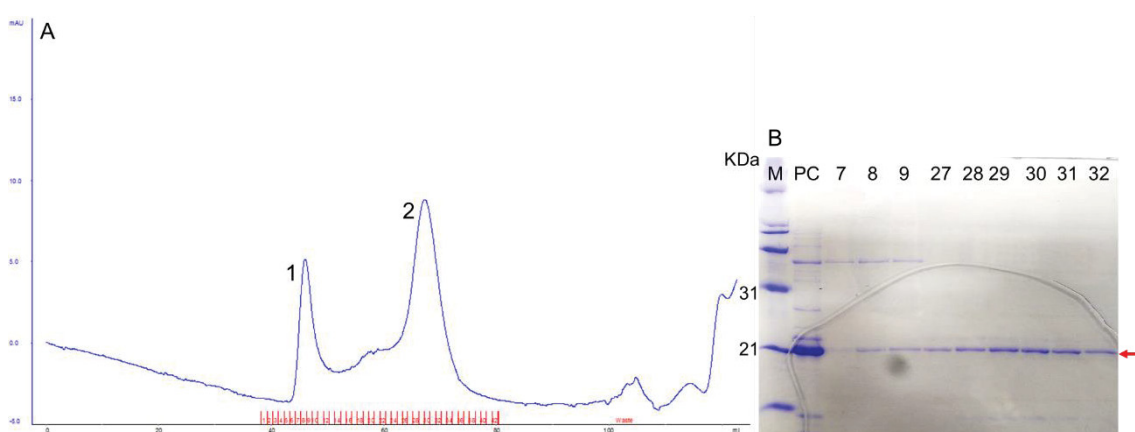


Figura 25 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E2 Δ S1-S2 por exclusão de tamanho em coluna Superdex 75 16/600. A) a linha azul corresponde à absorvância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna. A seta vermelha indica a proteína de interesse.

O pico 1 eluído em torno de 50 mL corresponde a um estado agregado da proteína, e o pico 2 eluído em torno de 60 – 70 mL corresponde ao estado

monomérico da proteína. O rendimento foi suficiente para a realização dos ensaios de interação por termoforese em microescala, a fração 32 foi concentrada e utilizada.

5.2.3 TcEIF4E3ΔN

A construção TcEIF4E3ΔN foi expressa a partir de 500 mL de cultivo bacteriano e purificado por gravidade, utilizando a resina NiSepharose 6 Fast Flow (GE Healthcare) (Figura 26).

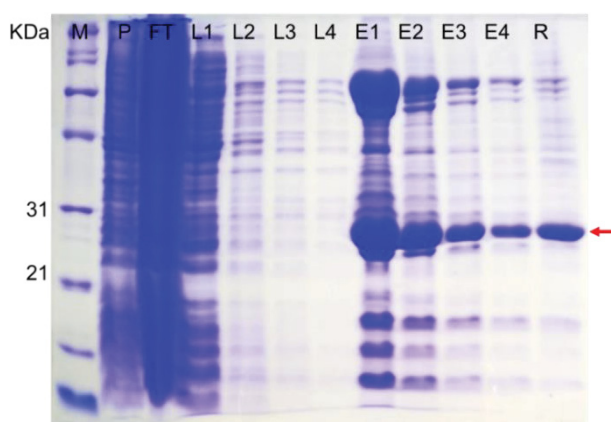


Figura 26 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E3ΔN. M: marcador de peso molecular (Bio-Rad Broad Range), P: pellet (sedimentado), FT: flowthrough, L1, L2, L3 e L4: lavagem 1, 2, 3 e 4 respectivamente, E1, E2, E3 e E4: eluição com tampão B 1, 2, 3 e 4 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

As frações E1, E2 e E3 foram reunidas e concentradas para a cromatografia de exclusão por tamanho.

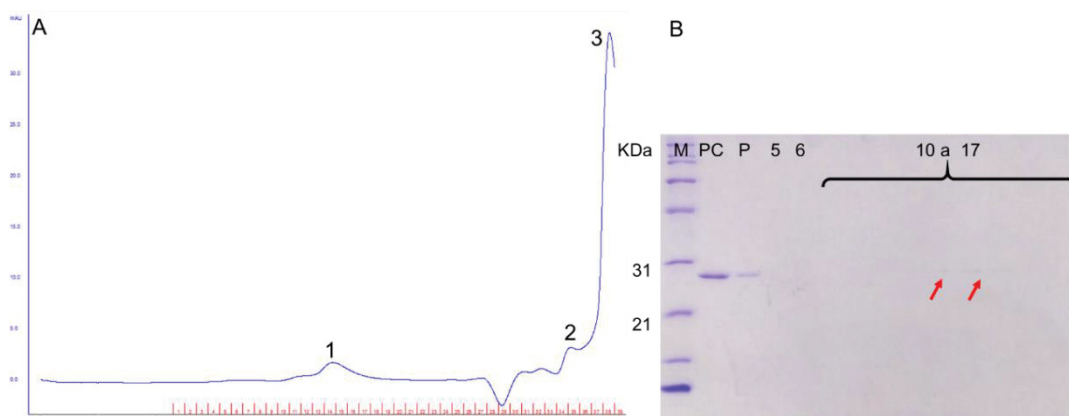


Figura 27 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E3ΔN por exclusão de tamanho em coluna Superdex 75 Increase 10/300. A) a linha azul corresponde à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; FT: flowthrough. A seta vermelha indica a proteína de interesse.

O resultado da cromatografia de exclusão por tamanho (Figura 27) não foi satisfatório, o pico 1 eluído em torno de 12 mL corresponde a proteína de interesse, porém a quantidade obtida nas frações 14 e 15 foram suficientes para a realização de um teste no ensaio de termoforese em microescala. A purificação foi repetida (resultado na Figura 28) e foi decidido aumentar o número de lavagens, bem como aumentar o número de eluições com o aumento de imidazol, afim de melhorar a pureza da proteína já na purificação de afinidade, uma vez que, na purificação por exclusão de tamanho há muita perda em quantidade de proteína. Acredita-se que a perda ocorra na concentração da amostra entre a purificação por gravidade e a exclusão por tamanho, a amostra possivelmente fica retida na membrana do concentrador o qual não conseguimos recuperar posteriormente.

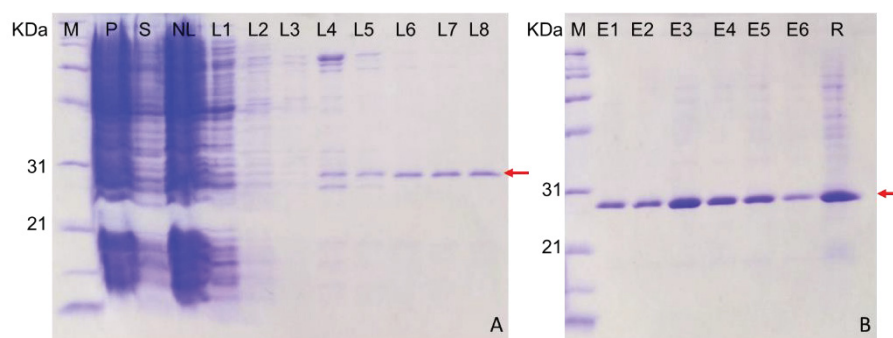


Figura 28 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E3ΔN. M: marcador de peso molecular (Bio-Rad Broad Range). A) P: pellet (sedimentado), S: fração solúvel. NL: flowthrough (fração não ligada), L1, L2, L3, L4, L5, L6, L7 e L8: lavagem 1, 2, 3, 4, 5, 6, 7 e 8 respectivamente. B) E1, E2, E3, E4, E5 e E6: eluição com tampão B 1, 2, 3, 4, 5 e 6 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

Com a mudança no número de lavagens e eluições foi obtido um grau de pureza suficiente para utilizar a proteína no ensaio de termoforese em microescala.

5.2.4 TcEIF4E4ΔN

A construção TcEIF4E4ΔN foi purificada a partir de 1 litro de cultura bacteriana. Iniciou pela purificação por gravidade, em resina NiSepharose 6 Fast Flow (GE Healthcare), e posteriormente as amostras foram submetidas à cromatografia de exclusão por tamanho.

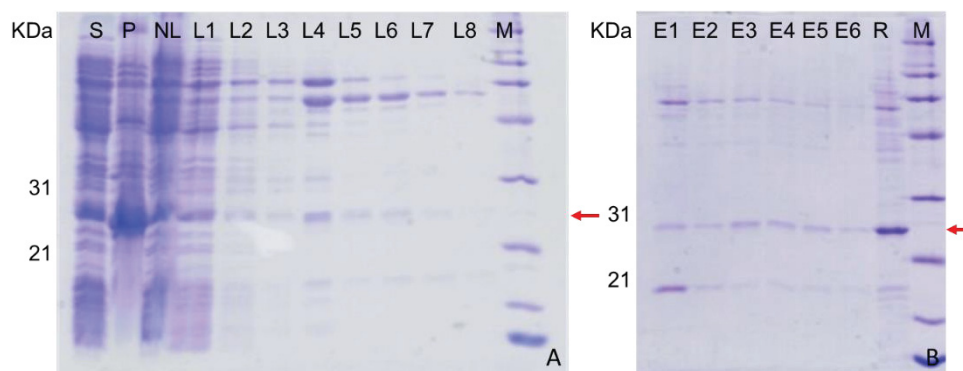


Figura 29 – Eletroforese em gel de poliácridamida da purificação de TcEIF4E4 Δ N. M: marcador de peso molecular (Bio-Rad Broad Range). A) S: fração solúvel; P: pellet (sedimentado); NL: flowthrough (fração não ligada), L1, L2, L3, L4, L5, L6, L7 e L8: lavagem 1, 2, 3, 4, 5, 6, 7 e 8 respectivamente. B) E1, E2, E3, E4, E5 e E6: eluição com tampão B 1, 2, 3, 4, 5 e 6 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

Para seguir para a cromatografia de exclusão por tamanho as frações da eluição 2 a 6 (cada fração possuía 300 μ L) foram reunidas e concentradas para um volume de 500 μ L.

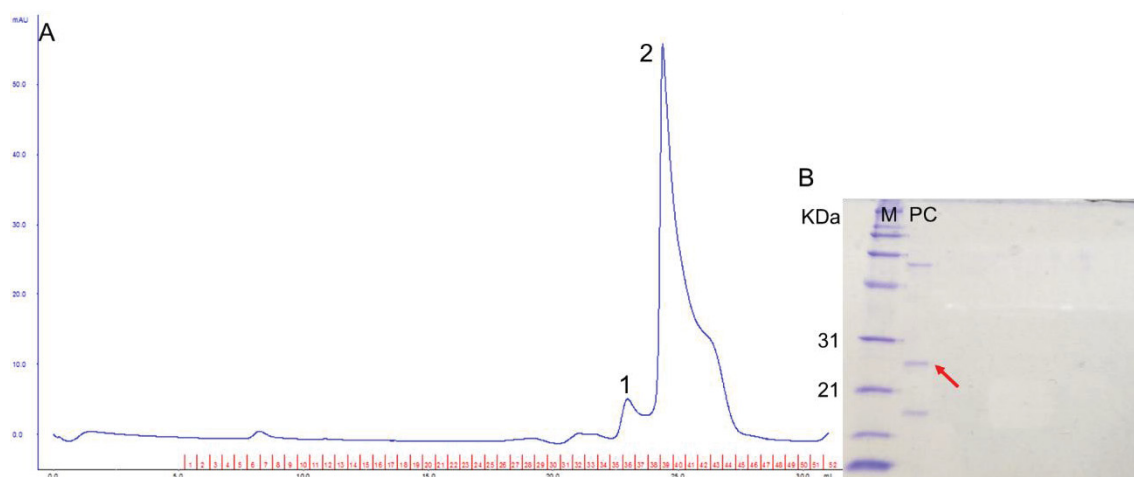


Figura 30 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E4 Δ N por exclusão de tamanho, em coluna Superdex 75 Increase 10/300. A) a linha azul corresponde à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna. A seta vermelha indica a proteína de interesse.

Acredita-se que parte da proteína fica retida na membrana do concentrador, assim como para TcEIF4E3 Δ N, é visível a perda da proteína na fração pré-coluna da eletroforese, a fração pré-coluna (amostra injetada na coluna) corresponde as frações da purificação por gravidade que foram reunidas e concentradas, a qual apresentou pouca quantidade da proteína de interesse (Figura 30).

5.2.5 TcEIF4E5 e TcEIF4E5ΔC

5.2.5.1 TcEIF4E5

A construção TcEIF4E5 foi expressa em volume de 500 mL de cultura bacteriana. A purificação constituiu da purificação por afinidade ao níquel e purificação de exclusão por tamanho.

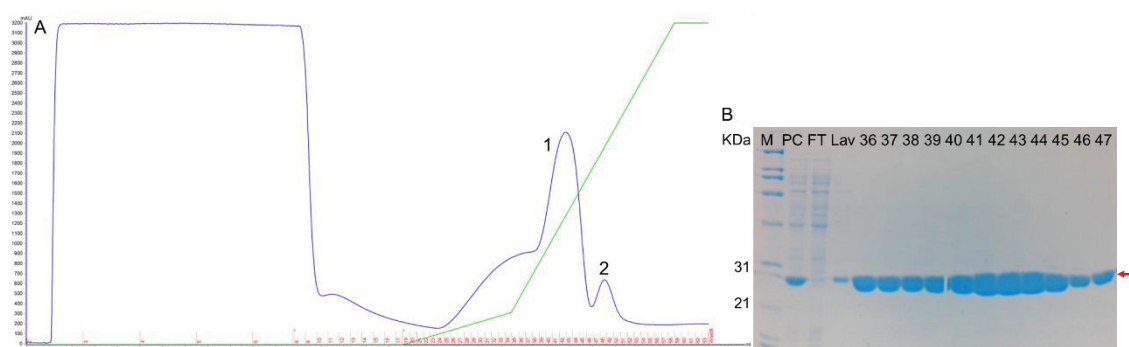


Figura 31 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E5 por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; FT: flowthrough; Lav: lavagem; 36 – 47: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

Todos os picos do cromatograma de afinidade (Figura 31) correspondem a proteína TcEIF4E5, porém apenas parte do pico 1 foi selecionado para a purificação de exclusão por tamanho. Foram selecionadas as frações 37 - 45 e concentradas para a cromatografia de exclusão por tamanho.

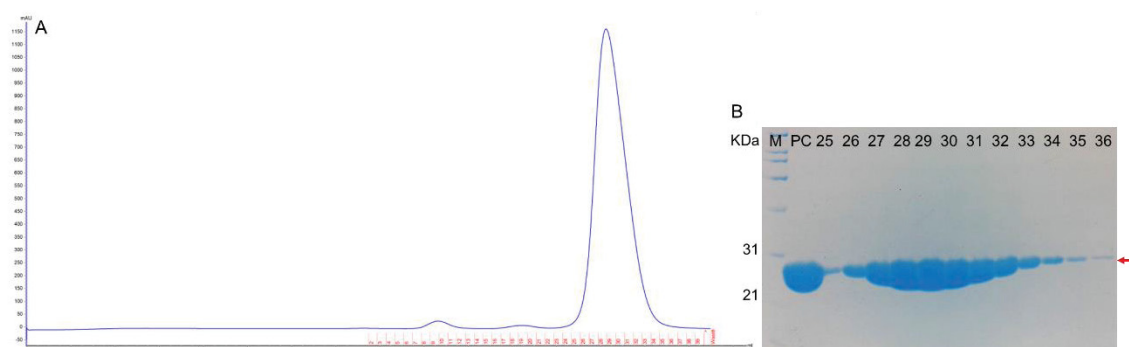


Figura 32 – Cromatograma (A) e eletroforese em gel de poliácridamida (B) da purificação de TcEIF4E5 por exclusão de tamanho, em coluna Superdex 75 16/600. A) a linha azul corresponde à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; 25 – 36: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A purificação da construção TcEIF4E5 foi satisfatória e rendeu proteína suficiente para seguir para os ensaios de interação por termoforese em microescala (foi utilizada a fração 33 para o ensaio) e para os ensaios de cristalização (foi concentrada a fração 30 para o ensaio).

5.2.5.2 TcEIF4E5 Δ C

A construção TcEIF4E5 Δ C foi purificada a partir de 500 mL de cultura bacteriana, foi realizado cromatografia de afinidade a níquel e cromatografia de exclusão por tamanho.

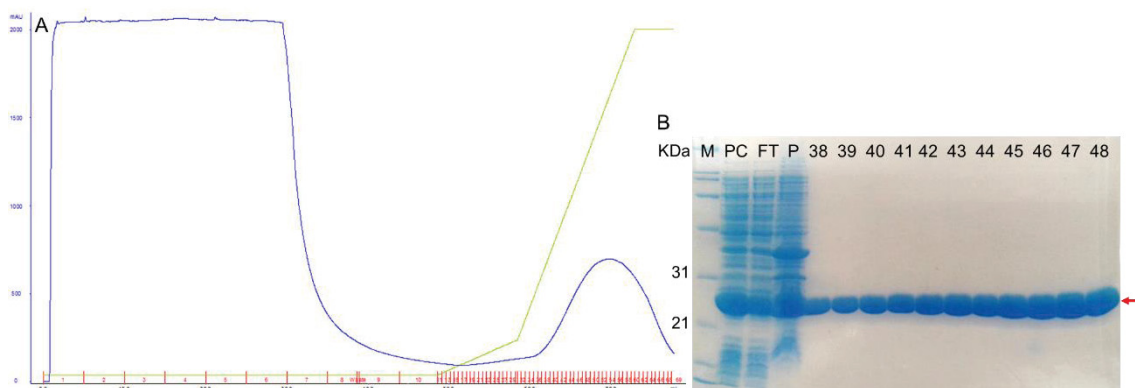


Figura 33 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E5 Δ C por afinidade a níquel. A) a linha em verde corresponde ao gradiente de tampão B e a linha azul à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; FT: flowthrough; P: pellet (sedimentado), 38 – 48: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A partir da cromatografia de afinidade a níquel foram selecionadas as frações 40 - 54 e concentradas para a cromatografia de exclusão por tamanho.

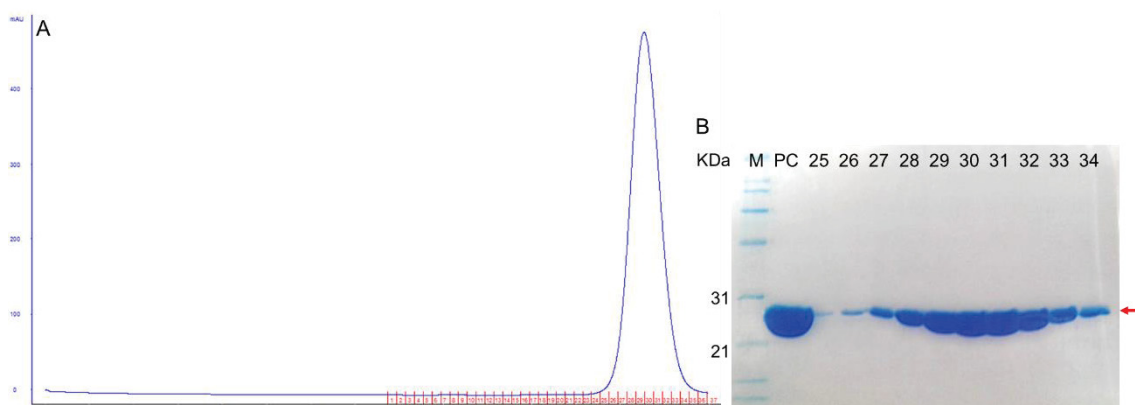


Figura 34 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E5 Δ C por exclusão de tamanho, em coluna Superdex 75 16/600. A) a linha azul corresponde à absorbância de UV de 280 nm. B) M: marcador de peso molecular (Bio-Rad Broad Range); PC: amostra pré-coluna; 25 – 34: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

O volume de eluição (entre 60 e 70 mL) do pico majoritário da cromatografia de exclusão por tamanho corresponde ao estado monomérico da proteína. O rendimento da purificação foi suficiente para o ensaio de interação por termoforese em microescala, o qual foi utilizado a fração 33 da gel filtração.

5.2.6 TcEIF4E6

A construção TcEIF4E6 foi purificada por gravidade. 500 mL de cultura bacteriana foi utilizada.

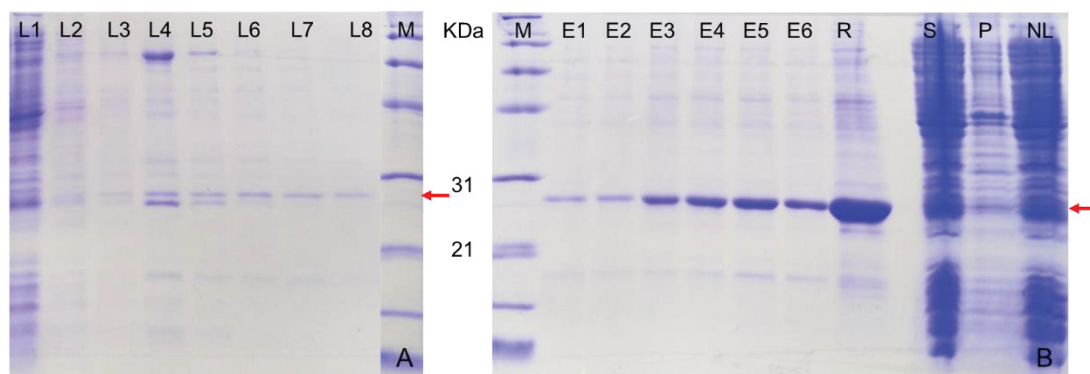


Figura 35 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E6. M: marcador de peso molecular (Bio-Rad Broad Range). A) L1, L2, L3, L4, L5, L6, L7 e L8: lavagem 1, 2, 3, 4, 5, 6, 7 e 8 respectivamente. B) E1, E2, E3, E4, E5 e E6: eluição com tampão B 1, 2, 3, 4, 5 e 6 respectivamente, R: resina; S: fração solúvel; P: pellet (sedimentado) e NL: flowthrough (fração não ligada). A seta vermelha indica a proteína de interesse.

As eluições 3, 4 e 5 foram reunidas e concentradas para a cromatografia de exclusão por tamanho.

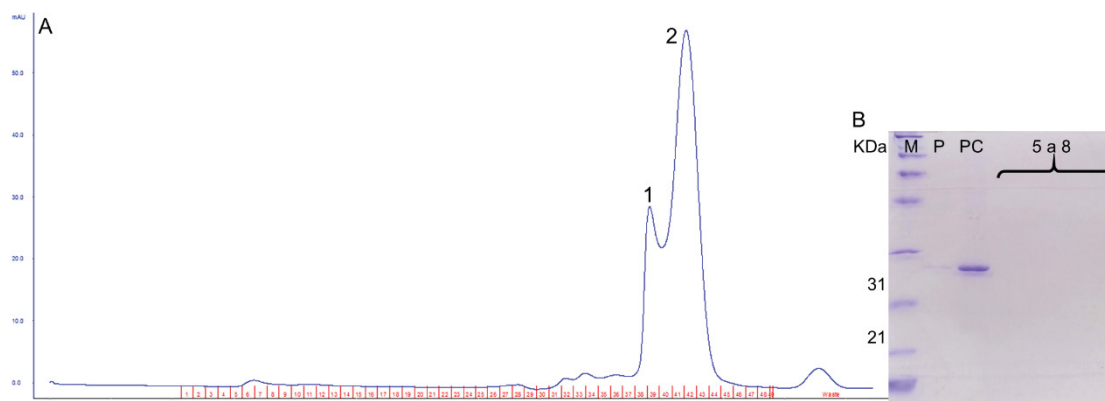


Figura 36 – Cromatograma (A) e eletroforese em gel de poliacrilamida (B) da purificação de TcEIF4E6 por exclusão de tamanho, em coluna Superdex 75 Increase 10/300. A) a linha azul corresponde à absorbância de UV de 280 nM. B) M: marcador de peso molecular (Bio-Rad Broad Range); P: pellet (sedimentado); PC: amostra pré-coluna; 5 a 8: frações correspondentes ao pico da eluição da proteína de interesse. A seta vermelha indica a proteína de interesse.

A proteína não foi eluída na cromatografia, o pico correspondente a proteína no estado monomérico deveria ter sido eluído em torno de 12 mL, os picos 1 e 2 correspondem a contaminantes da amostra de 1 KDa. Foi utilizado para o ensaio de interação a amostra pré-coluna. A purificação e o ensaio de interação foram repetidos (Figura 37).

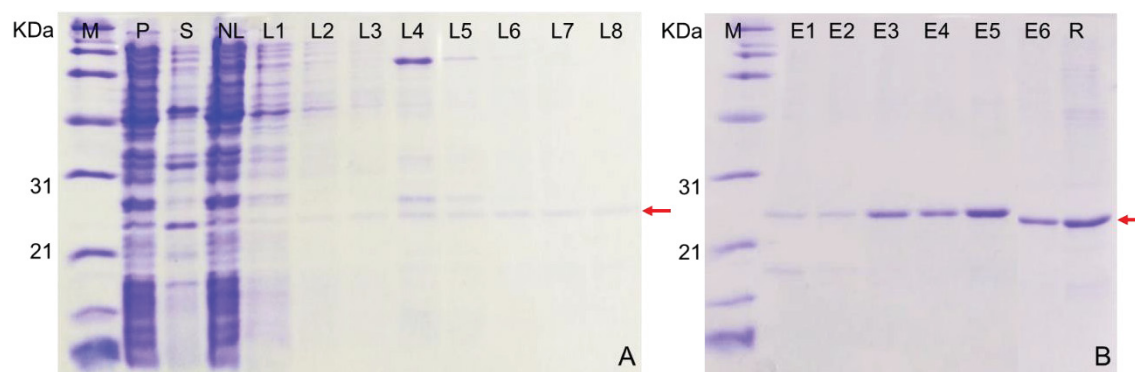


Figura 37 – Eletroforese em gel de poliacrilamida da purificação de TcEIF4E6. M: marcador de peso molecular (Bio-Rad Broad Range). A) P: pellet (sedimentado); S: fração solúvel; NL: flowthrough (fração não ligada), L1, L2, L3, L4, L5, L6, L7 e L8: lavagem 1, 2, 3, 4, 5, 6, 7 e 8 respectivamente. B) E1, E2, E3, E4, E5 e E6: eluição com tampão B 1, 2, 3, 4, 5 e 6 respectivamente, R: resina. A seta vermelha indica a proteína de interesse.

Para o segundo ensaio de interação a amostra utilizada foi a fração E6 logo após a purificação de afinidade. Foi utilizado a coluna PD SpinTrap G-25 (GE Healthcare) para a troca do tampão da amostra para a realização do ensaio de interação.

5.3 TERMOFORESE EM MICROESCALA

As proteínas purificadas foram utilizadas nos experimentos de interação com os análogos do cap por termoforese em microescala. Após a última etapa de purificação as frações foram quantificadas e selecionadas. Para a marcação fluorescente utilizou-se o kit RED-tris-NTA (NanoTemper) com a sonda NT 647, para marcação não covalente da cauda de hexahistidinas, seguindo protocolo do fabricante.

5.3.1 TcEIF4E1, TcEIF4E1 Δ N e TcEIF4E1 Δ S4-H2

A concentração de proteína utilizada nos ensaios foi de 12,5 nM. Os análogos do cap foram utilizados em uma concentração máxima de 500 μ M. TcEIF4E1 foi testada com dois análogos do cap (m^7 GTP e cap-4). A Figura 38 mostra os resultados dos experimentos de termoforese, expressos pela diferença de fluorescência normalizada (ΔF_{Norm}) em função da concentração do ligante. Caso ocorra a interação entre a proteína e o ligante a curva apresentará uma diferença na fluorescência normalizada (eixo Y do gráfico) conforme o aumento da concentração do ligante. Em vermelho e laranja são as curvas de TcEIF4E1 com cap-4 e m^7 GTP respectivamente. Em roxo e lilás as curvas de TcEIF4E1 Δ N com cap-4 e m^7 GTP respectivamente. E em verde escuro e verde claro as curvas de TcEIF4E1 Δ S4-H2 com cap-4 e m^7 GTP respectivamente.

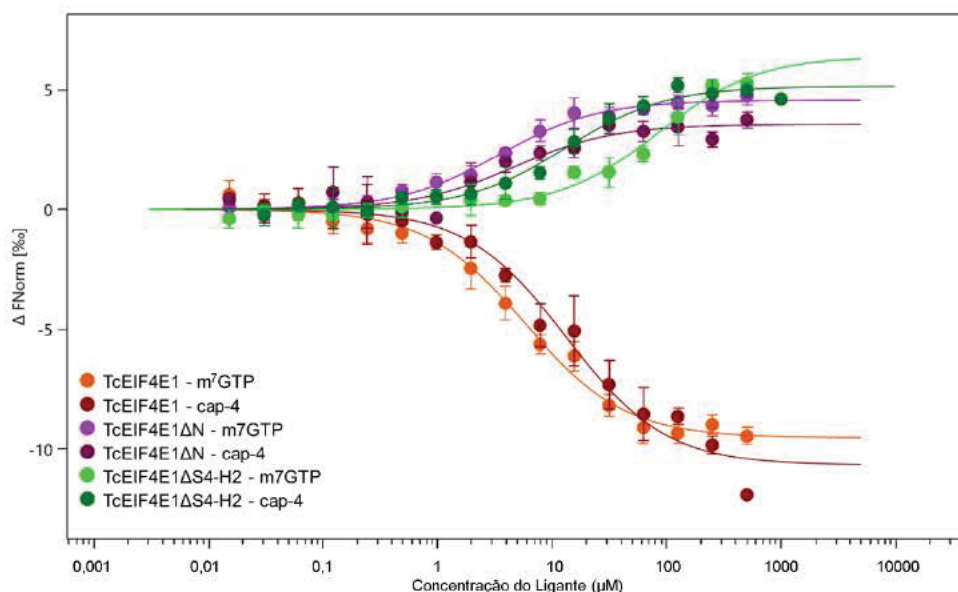


Figura 38 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E1. O gráfico representa a diferença de fluorescência normalizada em função da concentração do ligante.

Os valores da constante de dissociação calculados a partir dos ensaios de termoforese em microescala com TcEIF4E1 se encontram na Tabela 10.

Construção	Kd m ⁷ GTP (μM)	Kd cap-4 (μM)
TcEIF4E1	5,6 ± 0,6	13 ± 2
TcEIF4E1ΔN	3,4 ± 0,3	4 ± 1
TcEIF4E1ΔS4-H2	84 ± 18	13 ± 2

Tabela 10 - Valores de constante de dissociação obtidos para as construções TcEIF4E1, TcEIF4E1ΔN e TcEIF4E1ΔS4-H2 testados com os análogos m⁷GTP e cap-4 por termoforese em microescala.

Em comparação com resultados descritos na literatura tais valores de constante de dissociação são concordantes com os obtidos para EIF4E1 de *L. major* (ver Tabela 2) (YOFFE et al., 2006). A construção TcEIF4E1ΔN foi a construção que se mostrou mais afim aos análogos do cap. A interação de LmEIF4E1 com o análogo m⁷GTP é regulada por 4E1-IP1, a região N-terminal é onde ocorre a interação fator EIF4E1 de *L. major* com o repressor 4E-IP1, indicando que a região N-terminal tem um papel na interação com os análogos do cap (Meleppattu et al., 2018). A construção TcEIF4E1ΔS4-H2 que possui a deleção dos resíduos 135 – 178 apresentou afinidade ao cap-4 semelhante à proteína inteira, mas um aumento na constante de dissociação ao m⁷GTP de mais de uma ordem de grandeza. Quando comparamos o modelo por homologia

de TcEIF4E1 com a estrutura cristalográfica de EIF4E1 de *L. major*, observa-se que a região S4-H2 em TcEIF4E1 corresponde em LmEIF4E1 a região de interação da proteína com o peptídeo inibitório 4E-IP1 (Figura 6) assim é provável que a deleção do loop S4-H2 em TcEIF4E1 não influencie diretamente na interação com os análogos do cap mas sim indiretamente por meio de interações com proteínas parceiras.

5.3.2 TcEIF4E2ΔN e TcEIF4E2ΔS1-S2

As construções TcEIF4E2ΔN e TcEIF4E2ΔS1-S2 foram testadas com os análogos m⁷GTP e cap-4, a concentração da proteína utilizada nos ensaios foi de 12,5 nM, e a maior concentração dos análogos do cap utilizado foi de 2,5 mM. A Figura 39 mostra o resultado da termoforese para ambas as proteínas. Em tons de verde as curvas de TcEIF4E2ΔN e em tons de azul as curvas de TcEIF4E2ΔS1-S2.

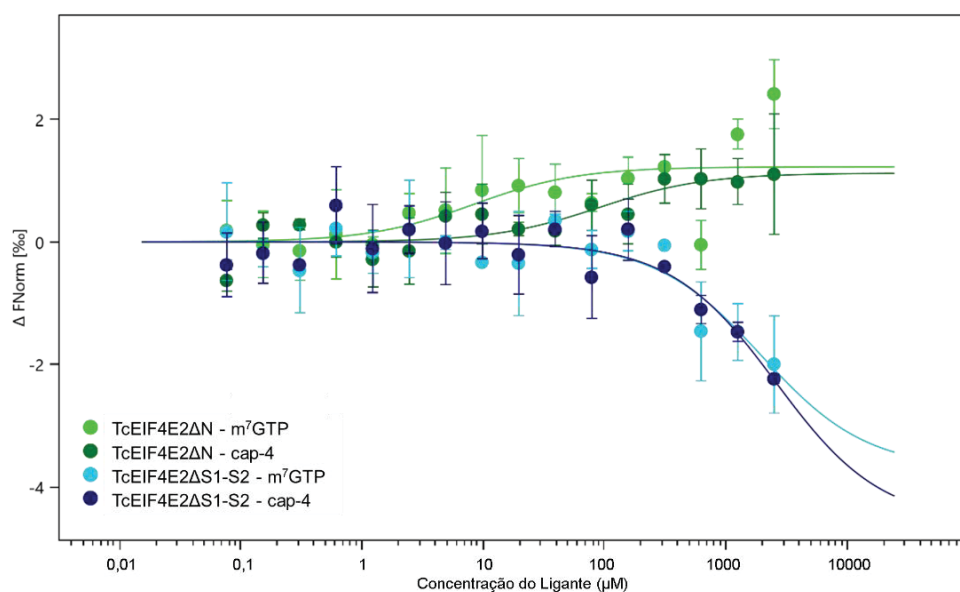


Figura 39 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E2. O gráfico representa a diferença de fluorescência normalizada em função da concentração do ligante.

Nenhuma das construções de TcEIF4E2 apresentou interação com os análogos nas condições do experimento, como constatado a partir da pequena variação da fluorescência normalizada em função da concentração do ligante. Tal resultado contraria os dados de que os homólogos 1 e 2 comporiam um

mesmo grupo dentre as homologas do fator 4E, os mesmos homólogos 1 e 2 em *L. major* possuem valores de interação altos e semelhantes ao homólogo 1, tanto para m⁷GTP como para cap-4 (ver Tabela 2).

O homólogo EIF4E2 em *T. cruzi* possui uma longa extensão de 37 resíduos (loop S1-S2) que conecta as fitas β 1- β 2, e esta região, segundo modelos estruturais (dados não mostrados), corresponderia a uma região muito próxima ao sítio de interação ao cap. Esta mesma região é mais extensa quando comparada a proteína humana (loop S1-S2 possui 12 resíduos) e estudos estruturais e eIF4E2 de humano indicam uma participação da região na interação com os análogos do cap (Rosettani et al., 2007).

O papel específico do loop S1-S2 do fator EIF4E2 de *T. cruzi* na interação com os análogos do cap infelizmente não pode ser avaliado, uma vez que, não obtivemos valores de interação para nenhuma das construções de TcEIF4E2. Mas há indícios de que o loop S1-S2 em TcEIF4E2 tenha um papel regulatório na interação com os análogos do cap por meio de rearranjos estruturais.

5.3.3 TcEIF4E5 e TcEIF4E5 Δ C

As construções TcEIF4E5 e TcEIF4E5 Δ C foram testadas com os análogos m⁷GTP e cap-4, a concentração da proteína utilizada nos ensaios foi de 12,5 nM, e a maior concentração dos análogos do cap utilizado foi de 2,5 nM. A Figura 40 ilustra as curvas normalizadas da termoforese. Em tons de roxo as curvas de TcEIF4E5 e em tons de vermelho as curvas de TcEIF4E5 Δ C.

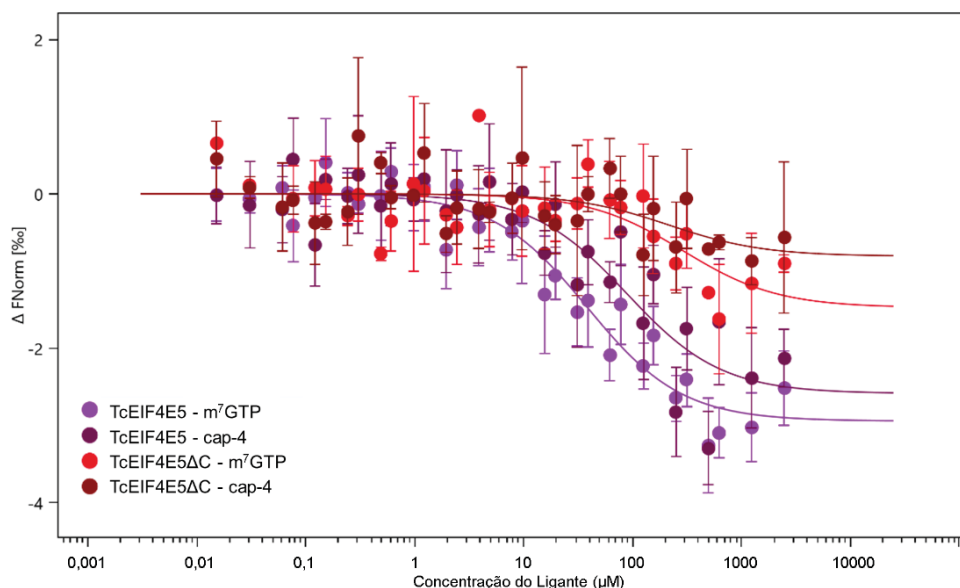


Figura 40 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E5. O gráfico representa a diferença de fluorescência normalizada em função da concentração do ligante.

Dentre as construções de TcEIF4E5 apenas a proteína inteira apresentou interação (K_d : $21 \pm 7 \mu\text{M}$ para interação TcEIF4E5- $m^7\text{GTP}$; e K_d : $110 \pm 33 \mu\text{M}$ para interação TcEIF4E5-cap-4), o que indica que a região C-terminal da proteína colabora com a interação do fator EIF4E ao cap (ver discussão no artigo em Anexo A). Os valores de K_d obtidos para as construções do homólogo 5 contrapõem os dados da literatura, sendo que o homólogo de EIF4E5 em *T. brucei*, apresenta maior afinidade por ambos os ligantes, com constantes de dissociação de $1,5 \mu\text{M}$ para $m^7\text{GTP}$ e de $1,8 \mu\text{M}$ para cap-4 (FREIRE et al., 2014a).

As extensões S1-S2 e S5-S6 os quais conectam as fitas $\beta 1$ - $\beta 2$ e $\beta 5$ - $\beta 6$ respectivamente, participam diretamente na interação com o cap-4 (Reolon et al., 2019) indicando assim que estas regiões tem um papel importante na interação com os análogos do cap e suas variações influenciam nas diferenças entre os homólogos de eIF4E.

5.3.4 TcEIF4E3ΔN e TcEIF4E6

As construções TcEIF4E3ΔN e TcEIF4E6 foram testadas com os análogos $m^7\text{GTP}$ e cap-4, a concentração da proteína utilizada nos ensaios foi de $12,5 \text{ nM}$, e a maior concentração dos análogos do cap utilizado foi de $2,5$

mM. A Figura 41 ilustra as curvas normalizada da termoforese. Em tons de vermelho as curvas de TcEIF4E3ΔN e em tons de azul as curvas de TcEIF4E6.

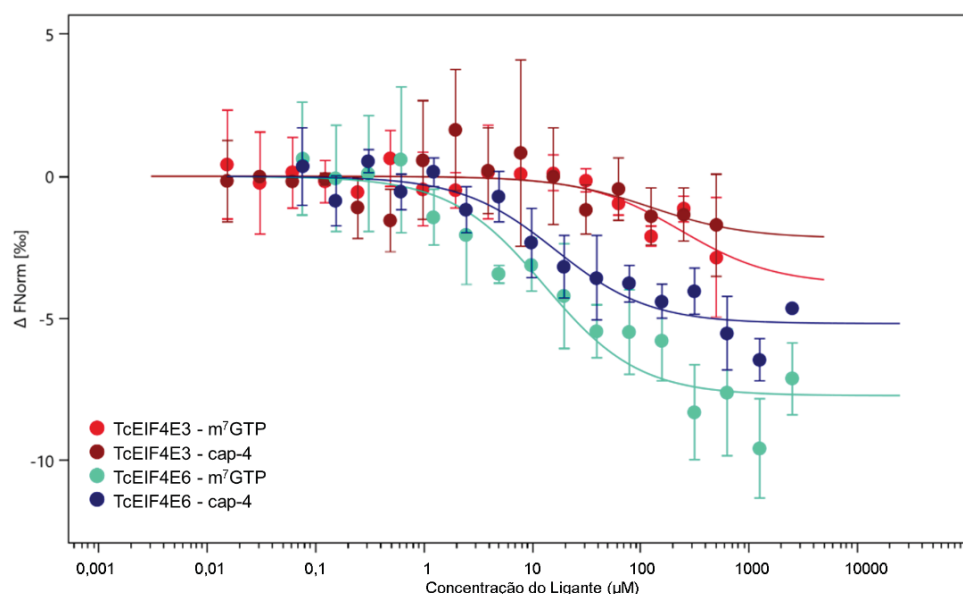


Figura 41 – Resultado dos experimentos de termoforese em microescala para as construções de TcEIF4E3 e TcEIF4E6. O gráfico representa a diferença de fluorescência normalizada em função da concentração do ligante.

Nas condições do nosso experimento detectou-se interação apenas do homólogo 6, TcEIF4E6-m⁷GTP com o valor de K_d estimado em $13 \pm 5 \mu\text{M}$, e TcEIF4E6-cap-4 com o valor de K_d estimado em $163 \pm 62 \mu\text{M}$. O homólogo TcEIF4E6 assim como TcEIF4E2 tem um loop S1-S2 mais extenso quando comparado a estrutura humana, em TcEIF4E6 compreendendo 18 resíduos.

Os os valores de interação com os análogos do cap obtidos para TcEIF4E6 corroboram com a hipótese de que a região do loop S1-S2 regula a interação com os análogos do cap e sua variabilidade implica em diferentes mecanismos moleculares entre os fatores EIF4E. Os dados obtidos são similares aos obtidos para o homólogo 5 de *T. cruzi* concordando com as propostas de que em tripanossomatídeos os homólogos 5 e 6 comporiam um mesmo grupo dentre os fatores EIF4E.

Surpreendentemente não detectamos interação com os análogos do cap pela construção de TcEIF4E3. Segundo os dados da literatura em *L. major* os homólogos 3 e 4 são contados como os mais prováveis de participarem do início da tradução, uma vez que, interagem com os fatores parceiros EIF4G (EIF4E3 interage com EIF4G4 e EIF4E4 interage com EIF4G3) (YOFFE et al., 2009), em

concordância em *T. brucei* foram detectados os mesmos complexos entre EIF4E3-EIF4G4 e EIF4E4-EIF4G3, além de que em *T. brucei* a proteína EIF4E3 é necessária para a sobrevivência do parasita, a sua depleção leva a um decréscimo na taxa de tradução (Freire et al., 2011). Assim a não detecção de interação do homólogo 3 em *T. cruzi* com os análogos do cap aqui testados indica que é possível que interação de EIF4E ao cap seja regulada por outras proteínas parceiras como o fator 4G, ou de que a função dos diferentes homólogos de eIF4E em *T. cruzi* não seja semelhante as outras espécies de tripanossomatídeos.

A Tabela 11 reúne os valores das constantes de dissociação determinadas para os homólogos de eIF4E de tripanossomatídeos.

Organismo	Construção	Kd (μ M) m ⁷ GTP	Kd (μ M) cap4	Referência
<i>T. cruzi</i>	TcEIF4E1	5,6 \pm 0,6	13 \pm 2	Este trabalho
<i>T. cruzi</i>	TcEIF4E1 Δ N	3,4 \pm 0,3	4 \pm 1	Este trabalho
<i>T. cruzi</i>	TcEIF4E1 Δ S4-H2	84 \pm 18	13 \pm 2	Este trabalho
<i>L. major</i>	LmIF4E1 (inteira)	5,98	4,35	Yoffe et al., 2006
<i>T. cruzi</i>	TcEIF4E2 Δ N	-	-	Este trabalho
<i>T. cruzi</i>	TcEIF4E2 Δ S1-S2	-	-	Este trabalho
<i>T. cruzi</i>	TcEIF4E3 Δ N	-	-	Este trabalho
<i>T. cruzi</i>	TcEIF4E5	21 \pm 7	110 \pm 33	Este trabalho (Reolon et al., 2019)
<i>T. brucei</i>	TbIF4E5 (inteira)	1,54	1,82	Freire et al., 2014a
<i>T. cruzi</i>	TcEIF4E6	13 \pm 5	163 \pm 62	Este trabalho
<i>T. brucei</i>	TbIF4E6 (inteira)	6,25	6,25	Freire et al., 2014b

Tabela 11 - Valores da constante de dissociação (Kd) para interação do fator EIF4E de tripanossomatídeos com análogos do cap.

É possível que a interação do fator EIF4E com os análogos do cap se modifique quando interagindo com um mRNA ou com outros parceiros, como o fator EIF4G e outras proteínas no complexo eIF4F (PRÉVÔT; DARLIX; OHLMANN, 2003). Importante também citar o papel da extensão N-terminal da proteína que em *T. cruzi* é predita como uma região desestruturada que pode ter a mobilidade para rearranjos estruturais, papel o qual não pode ser avaliado com

as proteínas produzidas neste trabalho os quais possuíam a deleção da região N-terminal.

5.4 ENSAIOS DE CRISTALIZAÇÃO

5.4.1 TcEIF4E2ΔN

Para o teste de cristalização de TcEIF4E2ΔN a proteína foi concentrada a 1,5 mg/mL, foram encontrados pequenos cristais (Figura 42) no kit PACT *premier* na condição D11 (0,2M cloreto de cálcio dihidratado, 0,1M Tris pH8 + 20% w/v PEG 6000). Não foi possível aumentar a concentração da proteína para novos testes pois a mesma precipitava. Novos kits foram testados, mas não foram obtidos melhores resultados.



Figura 42 - Cristais obtidos de TcEIF4E2ΔN. Cristais obtidos na condição de 0,2M cloreto de cálcio dihidratado + 0,1M Tris pH8 + 20% w/v PEG 6000.

5.4.2 TcEIF4E1ΔN

Obteve-se resultados promissores em ensaios preliminares de cristalização da proteína TcEIF4E1ΔN. Dos dois kits testados (JCSG+ e Index^{HT}) várias condições apresentaram precipitados cristalinos, dois refinamentos posteriores foram realizados.

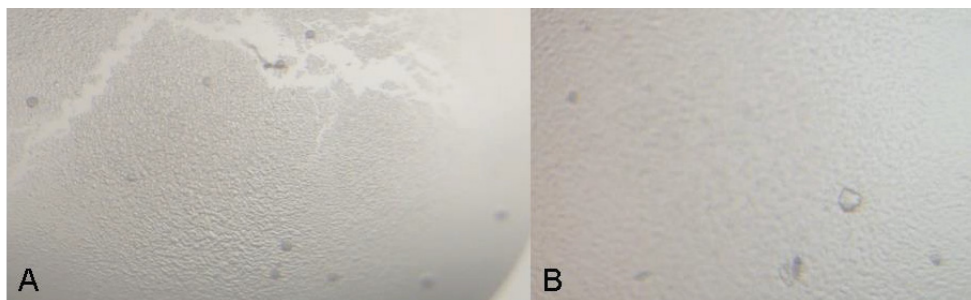


Figura 43 - Cristais obtidos de TcEIF4E1ΔN. A) cristais obtidos na condição 0,1M Bis-Tris pH 6,5 + 19% PEG 3350. B) cristais obtidos na condição 0,1M Bis-Tris pH 6 + 28% PEG 3350.

As melhores condições obtidas estão na Figura 43, em A) a condição pertence ao primeiro refinamento (0,1M Bis-Tris pH 6,5 + 19% PEG 3350, concentração da proteína 5 mg/mL). E em B) a condição pertence ao segundo refinamento (0,1M Bis-Tris pH 6 + 28% PEG 3350, com a proteína a 4 mg/mL).

5.4.3 TcEIF4E5

Obtiveram-se resultados promissores em ensaios preliminares de cristalização da proteína TcEIF4E5. Dos três kits testados (JCSG+, PACT *premier* e Index^{HT}) duas condições, H4 do kit Index^{HT} (0,2 M citrato de amônia pH 7 + 20% w/v PEG 3350) e D11 no kit PACT *premier* (0,2 M cloreto de magnésio hexahidratado + 0,1 M tris HCl pH 8 + 20% w/v PEG 6000) apresentaram os respectivos cristais da Figura 44.

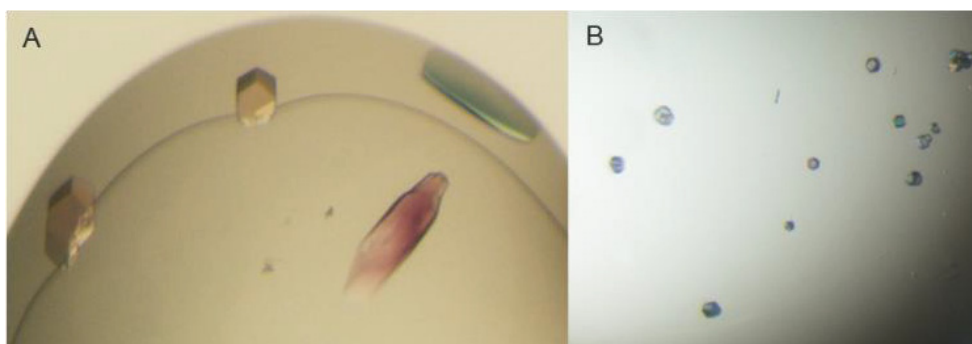


Figura 44 - Cristais obtidos em testes preliminares de TcEIF4E5. A) cristais obtidos na condição 0,2 M citrato de amônia pH 7 + 20% w/v PEG 3350. B) cristais obtidos na condição 0,2 M cloreto de magnésio hexahidratado + 0,1 M tris HCl pH 8 + 20% w/v PEG 6000.

A proteína TcEIF4E5 foi também produzida em meio mínimo M9 para marcação com selênio-metionina (a expressão e purificação foi realizada pela colaboradora Tatiana Assunção, o protocolo está no Anexo B), para posterior determinação da estrutura cristalográfica por métodos de difração anômala.

A proteína marcada foi testada nos mesmos refinamentos já mencionados para a proteína TcEIF4E5 não marcada, bem como aplicada nos kits JCSG+ e Index^{HT}. Em duas condições foram encontrados cristais, conforme a Figura 45.

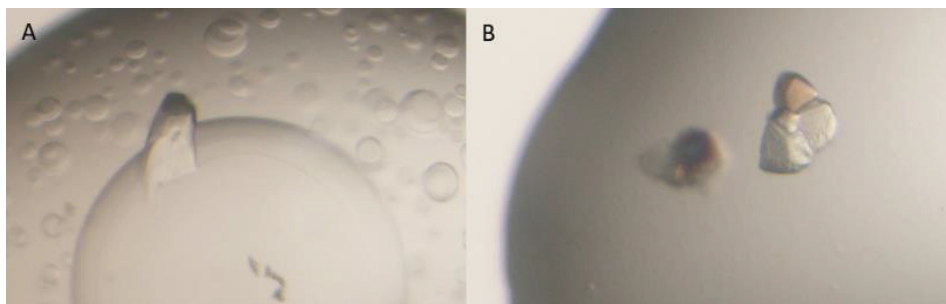


Figura 45 - Cristais obtidos de TcEIF4E5 marcados com selênio-metionina. A) cristais obtidos na condição 0,2 M de Sulfato de Amônio + 0,1 M HEPES pH 7,5 + 25% PEG 3350. B) cristais obtidos na condição 2 M de Sulfato de Amônio + 0,1 M Bis-Tris pH 5,5.

Para os refinamentos onde a proteína foi incubada com os análogos do cap (cap-1 e cap-4), as condições 0,1 M Bis-Tris pH 6 + 30% PEG 3350; 0,1 M Bis-Tris pH 6 + 32% PEG 3350 apresentaram cristais dos complexos TcEIF4E5-cap1 e TcEIF4E5-cap4 respectivamente (Figura 46). A condições testadas para as proteínas em complexo com os análogos do cap derivaram de tentativas testadas anteriormente no nosso laboratório, os quais foram obtidos os cristais do complexo TcEIF4E5-m⁷GTP na condição 0,1 M Bis-Tris pH 6 + 0,2 M Sulfato de Lítio + 25% PEG 3350. (MATOS, B. 2018).

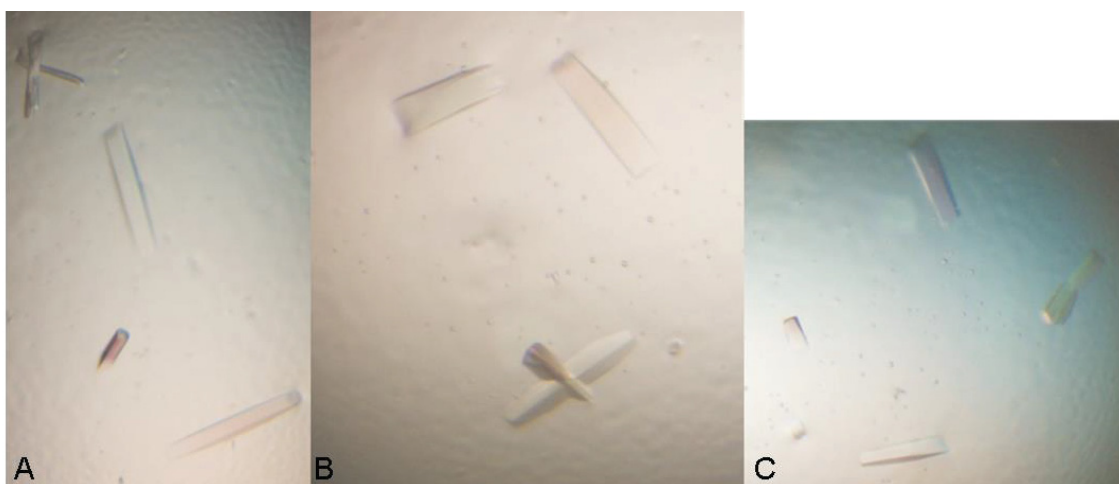


Figura 46 - Cristais obtidos de TcEIF4E5 com análogos do cap. A) Cristais do complexo TcEIF4E5-cap1 obtidos na condição 0,1 M Bis-Tris pH 6 + 30% PEG 3350. B) e C) Cristais do complexo TcEIF4E5-cap-4 obtidos na condição 0,1 M Bis-Tris pH 6 + 32% PEG 3350.

5.5 DETERMINAÇÃO E ANÁLISE DA ESTRUTURA CRISTALOGRAFICA DE TcEIF4E5

Os cristais obtidos de TcEIF4E5 foram submetidos à coleta de dados de difração de raios X. A coleta de dados de difração foi realizada no Síncrotron SOLEIL (França). O conjunto de dados dos cristais de TcEIF4E5 em complexo com o m⁷GTP havia sido obtido anteriormente no nosso laboratório (MATOS, B. 2018). A determinação da estrutura cristalográfica de TcEIF4E5 foi por difração anômala utilizando a proteína marcada com selênio-metionina e determinação da estrutura dos complexos TcEIF4E5-m⁷GTP, TcEIF4E5-cap-1 e TcEIF4E5-cap-4 foi por substituição molecular. Os detalhes da determinação e análises da estrutura foram publicados no artigo: Crystal structure of the *Trypanosoma cruzi* EIF4E5 translation factor homologue in complex with mRNA cap-4, publicado na revista Nucleic Acids Research em 2019, no volume 47, número 11. O artigo se encontra no Anexo C.

Como esperado, a estrutura de TcEIF4E5 mostra o enovelamento característico ao da família eIF4E, com formato em C, onde na face interna está o sítio de interação ao cap. Foi determinado a estrutura de TcEIF4E5 em complexo com os análogos m⁷GTP, cap-1 e cap-4. Os análogos do cap possuem a densidade eletrônica bem definida, exceto pela base do segundo nucleotídeo de cap-1 (Figura 3 no artigo) A segunda ribose e base de cap-1 apresenta uma conformação diferente quando comparada com cap-4, provavelmente pela flexibilidade do análogo no complexo.

A estrutura do complexo TcEIF4E5-cap-4 revela um grande número de interações entre o ligante e a superfície da proteína. TcEIF4E5 conserva a interação por *stacking* com m⁷G dos análogos do cap, por meio dos resíduos aromáticos W33 e Y83 (equivalentes aos resíduos W56 e W102 de eIF4E1 humano), formando a interação do tipo sanduiche. As bases de cap-4 se mostram empacotadas formando uma conformação em *stacking*, orientado na direção da proteína deixando os grupos fosfatos expostos ao solvente, a metilação das bases de cap-4 se mostra importantes para o posicionamento do análogo na interação com o complexo TcEIF4E5-cap-4 (Figura 4 no artigo).

As principais diferenças de TcEIF4E5 com os outros membros da família são as regiões conectoras de $\beta 1$ - $\beta 2$, $\beta 3$ - $\beta 4$ e $\beta 5$ - $\beta 6$, os loops S1-S2, S3-S4 e S5-S6 respectivamente são mais extensos quando comparados a estrutura humana, estas regiões se destacam pois realizam várias interações com as bases de cap-4 (Figura 5 no artigo).

Os homólogos de eIF4E em *T. cruzi* de forma geral apresentam maiores extensões nas regiões conectoras de $\beta 1$ - $\beta 2$, $\beta 3$ - $\beta 4$ e $\beta 5$ - $\beta 6$, tais regiões se mostram importantes para a interação com os análogos do cap, uma vez que a flexibilidade dos loops permitem a proteína se ajustar a diferentes ligantes, como é visto nos complexos de TcEIF4E5 com m⁷GTP, cap-1 e cap-4.

A análise da superfície eletrostática de TcEIF4E5 mostra que especificamente a região de interação da proteína com os fosfatos do cap é positiva enquanto o restante da superfície da proteína é predominantemente neutra ou ligeiramente acídica (Figura 7A no artigo). O padrão da superfície eletrostática em TcEIF4E5 dita o modo de interação e a conformação de cap-4, com as bases do cap interagindo entre si e direcionadas para a proteína. É importante destacar a comparação entre a superfície eletrostática de TcEIF4E5 com a superfície de *L. major* EIF4E1 (Figura 7B no artigo), onde o loop S5-S6 em LmEIF4E1 apresenta um potencial eletrostático positivo, assim, contrário de TcEIF4E5, é possível que a conformação e o modo de interação de cap-4 com o EIF4E1 de *L. major* seja predominantemente ditado pelos fosfatos do cap, deixando as bases do mesmo expostas. Os valores de interação diferentes obtidos por cada homólogo e em cada espécie indica diferenças moleculares entre os fatores EIF4E e consequentemente funções diferentes para cada um.

6 CONSIDERAÇÕES FINAIS

O objetivo de purificar os 6 homólogos do fator eIF4E de *T. cruzi* foi alcançado exceto pelo fator EIF4E4, para o qual ainda é necessária a padronização na expressão e purificação. Com a padronização das purificações das construções de TcEIF4E foi possível realizar os testes de interação com os análogos do cap por termoforese em microescala.

A estrutura cristalográfica de TcEIF4E5 mostrou pela primeira vez a estrutura tridimensional de uma proteína da família em complexo com o cap-4, característico de tripanossomatídeos. A estrutura do complexo revelou uma conformação distinta da interação entre RNA-proteína não sendo via fosfatos, mostrando a importância das metilações extras de cap-4 em relação a m⁷GTP.

Os testes de termoforese em microescala mostraram que dentre os cinco fatores EIF4E de *T. cruzi* analisados, TcEIF4E1 é o que apresenta maior afinidade aos análogos do cap. A deleção da inserção S4-H2 em TcEIF4E1 influencia na interação com o análogo m⁷GTP, reduzindo a afinidade. A maior afinidade é observada para a construção sem a extensão N-terminal (TcEIF4E1ΔN), com valor comparável ao ortólogo em *L. major*. Do ponto de vista funcional, nossos resultados sugerem que TcEIF4E1 e TcEIF4E2 não pertencem a um mesmo grupo como proposto por Freire e colaboradores (2017), pois TcEIF4E2 não apresenta interação com nenhum análogo do cap. O EIF4E5 de *T. brucei* apresenta maior afinidade aos análogos em relação à EIF4E5 de *T. cruzi*, o mesmo acontece para EIF4E6. Porém, nossos resultados mostram que EIF4E5 e EIF4E6 de *T. cruzi* apresentam valores de interação aos análogos do cap muito semelhantes, o que concorda com o proposto por Freire (2017) de que ambos fariam parte de um mesmo grupo dentre os homólogos de eIF4E. Para TcEIF4E3 não observamos interação com os análogos do cap, o que diverge do que esperávamos, pois como descrito para *L. major* e *T. brucei* o fator EIF4E3, assim como EIF4E4, seriam os principais fatores que participam na formação de EIF4F e no processo de iniciação da tradução.

Os resultados dos ensaios de interação por termoforese em microescala indicam que os mecanismos de interação entre os fatores de iniciação EIF4E e o mRNA podem variar entre as espécies de tripanossomatídeos. Além disso, tal

mecanismo parece ser também específico para cada um dos seis fatores EIF4E. Assim, com os dados de interação aqui obtidos não podemos afirmar qual ou quais dos homólogos de eIF4E em *T. cruzi* teriam o papel de formar o complexo eIF4F e participar no início da tradução, pois apesar de TcEIF4E1 apresentar a mais alta afinidade pelos análogos do cap, ainda não foi relatada nenhuma interação do mesmo com um homólogo do fator eIF4G, componente do complexo eIF4F.

Estudos adicionais com os seis homólogos de eIF4E em *T. cruzi* ainda precisam ser realizados, a determinação da estrutura de TcEIF4E5 permitiu o entendimento da interação de TcEIF4E5-cap-4, porém o modo de interação entre os demais homólogos possui diferenças, como pode ser visto pelos valores diferentes nas constantes de dissociação obtidos no ensaio de termoforese em microescala. Estudos envolvendo outros parceiros de interação de EIF4E podem ser realizados, ensaios envolvendo o fator EIF4G e a proteína 4E-BP, pois possivelmente influenciem na interação com os análogos do cap. Além disso, a caracterização *in vivo* da função dos seis homólogos de eIF4E de *T. cruzi* no processo de tradução e/ou sua regulação é fundamental para uma melhor compreensão da importância dos diferentes fatores para a maquinaria de síntese proteica nestes parasitas.

7 REFERÊNCIAS

AITKEN, C. E.; LORSCH, J. R. A mechanistic overview of translation initiation in eukaryotes. *Nature structural & molecular biology*, v. 19, n. 6, p. 568–76, 2012.

ALBERTS et al. **Biologia molecular da célula**. 5. ed. Porto Alegre: Artmed, 2010.

BANGS, D. et al. Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. **Journal of Biological Chemistry**, v.267, p.9805-9815, 1992.

CAJAL, S., CASTELLVI, J., HUMMER, S., PEG, V., PELLETIER, J., SONENBERG, N. Beyond molecular tumor heterogeneity: protein synthesis takes control. **Oncogene**. v. 37, p. 2490–2501, 2018.

DE MELO NETO, O., DA COSTA LIMA, T., XAVIER, C., NASCIMENTO, L., ROMÃO, T., ASSIS, L., PEREIRA, M., REIS, C., PAPADOPOULOU, B. The unique Leishmania eIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. **RNA Biology**, v.12, p.1209-1221, 2015.

DHALIA, R., REIS, C., FREIRE, E., ROCHA, P., KATZ, R.; MUNIZ, J., STANDART, N., DE MELO NETO, O. Translation initiation in *Leishmania major*: Characterisation of multiple eIF4E subunit homologues. **Molecular and Biochemical Parasitology**, v.140, p.23-41, 2005.

FREIRE, et al. The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. **Molecular and Biochemical Parasitology**, v. 176, n. 1, p. 25–36, 2011.

aFREIRE, et al. eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in *Trypanosoma brucei*. **RNA Journal**. v. 20. n. 8, 2014.

bFREIRE, et al. *Trypanosoma brucei* translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog. **Eukaryotic Cell**. v. 13, n. 7, p. 896–908, 2014.

FREIRE, E., STURM, N., CAMPBELL, D., DE MELO NETO, O. The role of cytoplasmic mRNA cap-binding protein complexes in *Trypanosoma brucei* and other trypanosomatids. **Pathogens**, v.55, p.1-25, 2017.

GOMEZ, C., RAMIREZ, E., CALIXTO-GALVEZ, M., MEDEL, O., RODRIGUEZ, M. Regulation of Gene Expression in Protozoa Parasites. **Journal of Biomedicine and Biotechnology**. v. 10, 2010.

HINNEBUSCH, A. The Scanning Mechanism of Eukaryotic Translation Initiation. **Annual Review of Biochemistry**. v. 83. p. 779–812, 2014.

JACKSON, R.J.; HELLEN, C.U.T.; PESTOVA, T.V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Mol Cell Biol*, v.11, p.113-127, 2010.

JAHNE, E.; CARDOSO, D. S.; ABRANTES, M.; CAVALCANTI, F. Perfil epidemiológico dos portadores de doença de chagas: dos indicadores de risco ao processo de enfrentamento da doença. v. 24, n. 1, p. 41–46, 2016.

JOSHI, B., LEE, K., MAEDER, D., JAGUS, R. Phylogenetic analysis of eIF4E-family members. **BMC Evolutionary Biology**. v. 5, p. 48, 2005.

LAMA, D., et al. Water-bridge mediates recognition of mRNA cap in eIF4E. **Cell Press**. v. 25, p. 188-194, 2017.

LANA, M., MACHADO, E. Biology of *Trypanosoma cruzi* and Biological Diversity. **American Trypanosomiasis Chagas Disease**. v.1. p. 339–363, 2010.

LEHNINGER, A., NELSON, D., COX, M. **Princípios de Bioquímica de Lehninger**. 5. ed. Porto Alegre: Artmed, 2011.

LIU, W., et al. Structural Insights into Parasite eIF4E Binding Specificity for m⁷G and m^{2,2,7}G mRNA Caps. **Journal of Biological Chemistry**. v. 284. n. 45. 2009.

MARCOTRIGIANO, J., GINGRAS, A., SONENBERG, N., BURLEY, S. Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. **Cell**. v. 89, p. 951–961, 1997.

MATOS, B. M., **Caracterização estrutural dos fatores de iniciação da tradução EIF4E de *Trypanosoma cruzi* e análise de interações com a TcEIF4G5 e análogos do cap do mRNA**. 2018. Dissertação (Mestrado em Ciências-Bioquímica). Universidade Federal do Paraná. Curitiba. 2018.

MELEPPATTU, S., et al. Structural basis for LeishIF4E-1 modulation by na interacting protein in the human parasite Leishmania major. **Nucleic Acids Research**. v. 46, p. 3791–3801, 2018.

MONTERO, H., PÉREZ_GILérez-Gil, G., SAMPIERI, C. Eukaryotic initiation factor 4A (eIF4A) during viral infections. **Virus Genes**. v. 55, p. 267–273, 2019.

NIEDZWIECKA, A. et al. Biophysical Studies of eIF4E Cap-binding Protein: Recognition of mRNA 5' Cap Structure and Synthetic Fragments of eIF4G and 4E-BP1 Proteins. **Journal of Molecular Biology**. v. 319, p. 615-635, 2002.

OSBORNE, M., VOLPON, L., KORNBLATT, J., CULJKOCI-KRALJACIC, B., BAGUET, A., BORDEN, K. eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. **Proceedings of the National Academy of Sciences**, v. 110, p. 3877–3882, 2013.

REOLON, L., VICHIER-GUERRE, S., MATOS, B., DUGUÉ, L., ASSUNÇÃO, T., ZANCHIN, N., POCHET, S., GUIMARÃES, B. Crystal structure of the *Trypanosoma cruzi* EIF4E5 translation factor homologue in complex with mRNA cap-4. **Nucleic Acids Research**, v. 47, p. 1, 2019.

ROSETTANI, P., KNAPP, S., VISMARA, M., RUSCONI, L., CAMERON, A. Structures of the human eIF4E homologous protein, h4EHP, in its m7GTP-bound and unliganded forms. *Journal of Molecular Biology*. v. 368, p. 691–705, 2007.

SAFAEE, N. et al. Interdomain allostery promotes assembly of Poly(A) mRNA with PABP and eIF4G. **Molecular Cell**, v. 48, p. 375–386, 2012.

TOMOO, K., et al. Crystal structures of 7-methylguanosine 5' -triphosphate (m7GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. **Biochemical Society**, v. 362, p. 539–544, 2002.

TOMOO, K., et al. Structural basis for mRNA Cap-Binding regulation of eukaryotic initiation factor 4E by 4E-binding protein, studied by spectroscopic, X-ray crystal structural, and molecular dynamics simulation methods. **Biochimica et Biophysica Acta** n. 1753. p. 191– 208. 2005.

YOFFE et al. Cap-binding activity of an eIF4E homolog from *Leishmania*. **RNA**. v. 10. n. 11. 2004.

YOFFE et al. Binding Specificities and Potential Roles of Isoforms of Eukaryotic Initiation Factor 4E in *Leishmania*. **Eukaryotic Cell**. v. 5, n. 12, p. 1969–1979, 2006.

YOFFE et al. Evolutionary changes in the *Leishmania* eIF4F complex involve variations in the eIF4E–eIF4G interactions. **Nucleic Acids Research**, v. 37, n. 10, p. 3243–3253, 2009.

ZAHA, A.; FERREIRA, H.; PASSAGLIA, L. **Biologia molecular básica**. 5. ed. Porto Alegre: Artmed, 2014.

ZINOVIEV, A., LÉGER, M., WAGNER, G., SHAPIRA, M. A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. **Nucleic Acids Research**, v. 39, n. 19, p. 8404–8415, 2011.

ZINOVIEV, A., SHAPIRA, M. Evolutionary conservation and Diversification of the translation initiation apparatus in Trypanossomatids. **Comparative and Functional Genomics**. 2012.

ANEXO A – Protocolo de purificação por gravidade

Os ensaios de purificação por gravidade foram realizados pela Tecnologista Tatiana Assunção da Plataforma de Purificação de Proteínas do Instituto Carlos Chagas (FIOCRUZ-PR).

Purificação de bancada – NiSepharose 6Fast Flow (GE Healthcare)

Tampões utilizados:

- Tampão A: 50 mM Tris pH 8,0; 0,3 M NaCl; 20 mM imidazol; 10 mM β ME.
- Tampão B: 50 mM Tris pH 8,0; 0,3 M NaCl; 0,5 M imidazol; 10 mM β ME.

Método:

- Ressuspender o pellet com 5 mL de tampão A.
- Adicionar 150 μ g/mL de lisozima e manter por 1 h sob agitação em gelo.
- Sonicar em amplitude 30%, 10 ciclos por 10 segundos.
- Adicionar 100 μ L da resina e incubar por 2 h com a fração solúvel, manter a 4°C sob agitação leve.
- Lavar 3 vezes com tampão A (2 vezes com 1 mL e 1 lavagem com 0,5 mL).
- Eluir 3 vezes com 150 μ L de tampão B.

ANEXO B – Protocolo de expressão em meio mínimo M9, e purificação de TcEIF4E5-SeMet

O ensaio de expressão de TcEIF4E5 em meio mínimo M9 foi realizado pela Tecnologista Tatiana Assunção da Plataforma de Purificação de Proteínas do Instituto Carlos Chagas (FIOCRUZ-PR). O protocolo foi desenvolvido por Helmholtz Zentrum München, German Research Center for Environmental Health (2010).

Meio mineral M9

Soluções estoque:

- Solução salina M9 (10X):

Componentes	Concentração (g/L)	Dissolver os sais em 800 mL de água e ajustar para pH 7,2 com NaOH. Adicionar água para volume final de 1 L e autoclavar por 15 min a 121°C.
Na ₂ HPO ₄ – 2H ₂ O	75,2	
KH ₂ PO ₄	30	
NaCl	5	
NH ₄ Cl	5	

- Solução de elementos traços (100X):

Componentes	Concentração		Dissolver o EDTA em 800 mL de água e ajustar para pH 7,5 com NaOH. Adicionar os demais componentes e adicionar água para volume final de 1L. Esterilizar a solução com filtro 0,22 µm.
EDTA	5 g/L	13,4 mM	
FeCl ₃ – 6H ₂ O	0,83 g/L	3,1 mM	
ZnCl ₂	84 mg/L	0,62 mM	
CuCl ₂ – 2H ₂ O	13 mg/L	76 µg	
CoCl ₂ – 2H ₂ O	10 mg/L	42 µg	
H ₃ BO ₃	10 mg/L	162 µg	
MnCl ₂ – 4H ₂ O	1,6 mg/L	8,1 µg	

Para 1 litro de meio mineral M9 adicione à 867 mL de água estéril:

Volume (mL)	Concentração do estoque	Componentes do estoque	Concentração final na solução
100	Solução salina M9 (10X)	$\text{Na}_2\text{HPO}_4 - 2\text{H}_2\text{O}$	33,7 mM
		KH_2PO_4	22 mM
		NaCl	8,55 mM
		NH_4Cl	9,35 mM
20	20% Glucose	Glucose	0,4%
1	1 M MgSO_4	MgSO_4	1 mM
0,3	1 M CaCl_2	CaCl_2	0,3 mM
1	Biotina (1mg/mL)	Biotina	1 μg
1	Tiamina (1mg/mL)	Tiamina	1 μg
10	Solução de elementos traços (100X)	Elementos traços	1X

*As soluções estoque devem estar autoclavadas (a biotina e a tiamina devem ser esterilizadas por filtração em membrana de 0,22 μM).

**ANEXO C - Artigo Crystal structure of the *Trypanosoma cruzi*
EIF4E5 translation factor homologue in complex with mRNA cap-4**

Crystal structure of the *Trypanosoma cruzi* EIF4E5 translation factor homologue in complex with mRNA cap-4

Lidia Watanabe Reolon^{1,2}, Sophie Vichier-Guerre³, Bruno Moisés de Matos^{1,2}, Laurence Dugué³, Tatiana Reichert da Silva Assunção¹, Nilson Ivo Tonin Zanchin¹, Sylvie Pochet^{3,*} and Beatriz Gomes Guimarães^{1,2,*}

¹Carlos Chagas Institute, Oswaldo Cruz Foundation, FIOCRUZ-PR, Curitiba, Paraná 81350-010, Brazil,

²Biochemistry Postgraduate Program, Federal University of Parana, Curitiba, Paraná 81530-000, Brazil and ³Unité de Chimie et Biocatalyse, Institut Pasteur, UMR3523 CNRS, Paris 75015, France

Received March 16, 2019; Revised April 18, 2019; Editorial Decision April 23, 2019; Accepted April 24, 2019

ABSTRACT

Association of the initiation factor eIF4E with the mRNA cap structure is a key step for translation. Trypanosomatids present six eIF4E homologues, showing a low conservation and also differing significantly from the IF4Es of multicellular eukaryotes. On the mRNA side, while in most eukaryotes the mRNA contains cap-0 (7-methyl-GTP), the trypanosomatid mRNA features a cap-4, which is formed by a cap-0, followed by the AACU sequence containing 2'-O-ribose methylations and base methylations on nucleotides 1 and 4. The studies on eIF4E-cap-4 interaction have been hindered by the difficulty to synthesize this rather elaborated cap-4 sequence. To overcome this problem, we applied a liquid-phase oligonucleotide synthesis strategy and describe for the first time the crystal structure of a trypanosomatid eIF4E (*T. cruzi* EIF4E5) in complex with cap-4.

The TcEIF4E5-cap-4 structure allowed a detailed description of the binding mechanism, revealing the interaction mode for the AACU sequence, with the bases packed in a parallel stacking conformation and involved, together with the methyl groups, in hydrophobic contacts with the protein. This binding mechanism evidences a distinct cap interaction mode in comparison with previously described eIF4E structures and may account for the difference of TcEIF4E5-cap-4 dissociation constant in comparison with other eIF4E homologues.

INTRODUCTION

Translation initiation in eukaryotes is a complex and highly regulated process dependent on the action of *trans* acting factors to assemble the large (60S) and small (40S) ribosomal subunits with an initiator methionyl-tRNA positioned over the start codon of the mRNA. Initiation begins with mRNA recruitment, through the association of the so-called eIF4F complex with the cap structure at the 5' end and binding of the poly(A)-binding protein (PABP) to the 3' end of the mRNA. The eIF4F complex comprises a scaffold and regulatory subunit, eIF4G, which interacts with the cap-binding factor eIF4E and the RNA helicase eIF4A. The eIF4G subunit interacts also with PABP forming the activated mRNP complex (reviewed in 1). Thus, the mRNA cap recognition by eIF4E is a key step for translation and any dysregulation of cap-binding proteins leads to numerous physiological consequences. Due to its critical role in cancer cell proliferation, eIF4E has also been pointed out as a potential therapeutic target (reviewed in 2).

The first three-dimensional structure described for the eukaryotic translation initiation factor eIF4E (3) revealed a curved-shape α/β protein with the cap binding site located on the concave basal surface of the protein, and the nucleotide base sandwiched between two tryptophan residues. In the past years, X-ray crystal structures and NMR solution structures of eIF4E from several organisms, including human, were determined in complex with cap analogues, peptides derived from eIF4G as well as the inhibitory protein 4E-BP. These structural studies, together with functional assays, provided details on the cap binding mechanism (4–6), on the interaction of eIF4G and 4E-BP at the convex side of eIF4E and on the motifs modulating these interactions to regulate translation initiation (7–11).

*To whom correspondence should be addressed. Tel: +55 41 3316 3225; Fax: +55 41 33163267; Email: beatriz.guimaraes@fiocruz.br
Correspondence may also be addressed to Sylvie Pochet. Tel: +33 1 40613805; Fax: +33 1 45688805; Email: sylvie.pochet@pasteur.fr

In addition to the canonical eukaryotic translation initiation factor 4E (also called eIF4E1), two other classes of eIF4E have been identified and were then called eIF4E2, or 4E homologous proteins (4EHP), and eIF4E3 (12,13). Both are characterized by substitutions of conserved tryptophan residues, including one of the tryptophans involved in the m⁷G base sandwich responsible for cap binding. Structural and functional characterization of human eIF4E2 (14) and mouse eIF4E3 (15) showed that, despite residue substitutions, the cap binding pocket is preserved, and that its binding activity is functionally relevant. However, in comparison with the canonical eIF4E, eIF4Es from classes 2 and 3 both feature a much lower affinity for m⁷GTP cap. Accordingly, class 2 and 3 members of eIF4E have been implicated in distinct physiological roles, such as translation of specific mRNA subsets or translational repression (reviewed in 16,17).

Kinetoplastida, a class of early-diverging eukaryotes, which includes pathogenic *Trypanosoma* and *Leishmania* species, display key differences in their translation machinery and interaction patterns. The first major difference involves the mRNA 5'-cap structure. While in most eukaryotes the mRNAs feature a cap-0 component, the trypanosomatid mRNAs have a cap-4. This cap-4 is composed of cap-0, followed by a hypermodified AACU sequence with 2'-*O*-ribose methylations on all residues, and a base methylation on nucleotides 1 and 4 (18). Moreover, trypanosomatid genomes encode for six divergent homologues of eIF4E (EIF4E1 to EIF4E6), with a low degree of conservation among themselves (sequence identity below 22%, except for EIF4E5 and EIF4E6, with ~30% of sequence identity), and significantly different from their counterparts from multicellular eukaryotes. In addition, five homologues of eIF4G and two or three of PABP are found in *Trypanosoma* and *Leishmania* species. Interaction between these homologues to form eIF4F-like and activated mRNP complexes has been shown to be species-specific (reviewed in 19,20).

It is reasonable to assume that these divergent trypanosomatid EIF4E homologues have undergone structural changes to enable them to interact with the unusual cap-4. Phylogenetic analysis revealed that Trypanosomatids EIF4E cluster in three groups, E1–E2, E3–E4, E5–E6, none of them belonging to the metazoan classes 1, 2 and 3 of eIF4E proteins (21,22). Studies of trypanosomatid EIF4Es have mostly focused on *Trypanosoma brucei* and *Leishmania* species. EIF4E3 and EIF4E4, which were described as the most abundant, are essential for cell survival and can form canonical eIF4F complexes implicated in translation (23,24). On the other hand, although limited functional data are available for EIF4E from groups 1 and 3, studies have shown that these are unlikely to be involved in general translation (reviewed in 20). Group 3 members are not essential in *T. brucei*. However, knockdown of TbEIF4E5 and TbEIF4E6 caused, respectively, loss of cell motility and defects in motility and instability of flagellar attachment (25,26). TbEIF4E5 was found in two distinct complexes. One including TbEIF4G1, a 47.5-kDa protein containing conserved RNA-binding domains and either the regulatory protein 14–3–3 II or a 117.5-kDa protein with guanylyltransferase and methyltransferase domains. The second complex was formed by TbEIF4E5, TbEIF4G2, a 17.9-kDa

hypothetical protein and both 14–3–3 variants I and II (25). TbEIF4E6, was found in a complex with the TbEIF4G5 homologue and a hypothetical protein of 70.3 kDa, referred to as TbG5-IP, which contains a P-loop NTPase domain (26). These phenotypes and interaction patterns suggest possible roles for these factors in specific cellular pathways.

In vitro interaction assays have shown that *Leishmania* and *T. brucei* EIF4E homologues have different affinities for cap-4 (25–27). Again, these differences are probably related to their function either in translation or in regulatory processes. Moreover, such differences in affinity, suggest a distinct binding mode to the mRNA cap-4 for each of these EIF4E homologues. Further information on trypanosomatid EIF4E-cap-4 interactions are dependent on structure determination of EIF4E-cap-4 complexes. So far, the three-dimensional structure of *L. major* EIF4E1 has been the only reported for a trypanosomatid EIF4E factor and it was described in its apo form (28). This may be partly explained by the fact that trypanosomatid EIF4E homologues are not easily amenable to structural studies. To overcome this problem, we screened *T. cruzi* EIF4E homologues for crystallization and obtained useful crystals for TbEIF4E5. Equally important, the unavailability of mRNA caps in milligram amount greatly hampers any biological and structural studies of these cap-binding proteins. Accordingly, with the key help of a large-scale chemical synthesis, we could prepare a cap-1 dinucleoside derivative in milligram amount as well as the hypermodified cap-4 in similar amount. With these cap derivatives, we have then obtained the crystal structures of *T. cruzi* EIF4E5 in complex with m⁷GTP, cap-1 and cap-4. These structures provide unprecedented details on the interaction of an EIF4E homologue with the mRNA cap-4, specific to trypanosomatids. Their analysis combined with biophysical assays revealed significant differences in comparison with EIF4E homologues of other eukaryotes. The unique features mediating *T. cruzi* EIF4E5-cap-4 interaction thus provide a molecular basis for its distinct mRNA affinity, which may account for the different functions of trypanosomatid homologues in translation mechanisms.

MATERIALS AND METHODS

Chemical synthesis of cap-1 and cap-4

General methods. Chemicals and solvents were purchased from Acros Organics or Sigma-Aldrich and were used without further purification. Nucleosides were purchased from Carbosynth. Standard RNA phosphoramidites and reagents for RNA synthesis were purchased from Sigma-Aldrich or Glen Research. Chemical reactions were monitored by thin layer chromatography (TLC) on pre-coated Merck silica gel plates (60 F₂₅₄/0.2 mm), by LC-MS (Agilent 1260 series, Quadrupole 6130) or by HPLC (Agilent 1200 series) using a C18 reverse phase column. Purification by flash chromatography was performed using Merck silica gel 60 (0.063–0.200 mm). Analytical HPLC was carried out on an Agilent 1200 series system equipped with a diode-array detector using a C18 reverse phase column (Kromasil, 5 µm, 100 Å, 4.6 × 150 mm) at a flow rate of 1 ml.min^{−1} and a linear gradient of acetonitrile in 10 mM triethylammonium acetate buffer (TEAA) over 20 min. Purification by

HPLC was carried out on an Agilent 1100 Series system using a C18 reverse phase column (Kromasil, 5 μ m, 100 Å, 10 \times 250 mm) at a flow rate of 4.0 ml min⁻¹, and an isocratic or linear gradient of acetonitrile in 10 mM TEAA over 20 min. Yields refer to chromatographically and spectroscopically pure (>95%) compounds. ¹H NMR, ¹³C NMR and ³¹P NMR spectra were recorded on a Bruker Advance 400 system operating at 400.13, 100.62 and 168.98 MHz, respectively. Chemical shifts are given in ppm (δ) relative to the residual solvent peak, coupling constants (*J*) are reported in Hertz and standard abbreviations are used. Assignment of ¹H and ¹³C signals was performed by analysis of the correlated homonuclear ¹H, ¹H-COSY and heteronuclear ¹H, ¹³C-HMBC, ¹H, ¹³C-HSQC spectra. High-resolution mass spectra were recorded on a Waters Q-TOF micro MS instrument under electrospray ionization in positive ionization mode using a mobile phase of acetonitrile/water with 0.1% formic acid, or negative ionization mode using a mobile phase of acetonitrile/water with 10 mM ammonium acetate. MALDI-MS analysis was performed on a Bruker UltrafleXtreme MALDI TOF/TOF in positive mode. Samples were dissolved in pure dichloromethane (1 mg/ml). An aliquot (1 μ l) was deposited onto the Maldi plate and the matrix solution (1 μ l, 10 mg/ml of 2,5-dihydroxybenzoic acid in 1:1 acetonitrile/water + 0.1% TFA) was immediately added. All mass spectra were externally calibrated from 700 to 3500 Da using the peptide calibration standard from Bruker Starter kit.

General procedures for RNA assembly.

Coupling and oxidation reaction. A solution of benzylthio-1*H*-tetrazole (BMT) (12.5 eq) in a mixture of dry acetonitrile (25 ml, 0.5 mmol) and dry dichloromethane (37.5 ml, 0.5 mmol) was added dropwise during 20 min to a premixed solution of amidite monomer (2.5–3.5 eq) and RNA_{*x*}-ACSS (0.5 mmol) in dry dichloromethane (37.5 ml) containing 4 Å molecular sieves (0.2 g) kept at 35°C. Reaction completion was confirmed by monitoring substrate consumption using TLC. After stirring for 1 h at room temperature, 2-butanone peroxide solution (BPO, 50 ml of 0.67% Luperox® in dichloromethane) was added dropwise, and the mixture was stirred at room temperature for 5 min. The resulting solution was diluted with methanol and concentrated under vacuum. After addition of methanol and filtration, the product DMT-RNA_{*x+1*}-ACSS was obtained as a white solid.

Deprotection. To DMT-RNA_{*x+1*}-ACSS in dichloromethane (53 ml, 0.5 mmol) was added 5% DCA in dichloromethane (17 ml). After stirring for 15 min at room temperature, the reaction mixture was diluted with methanol (38 ml) and concentrated under vacuum. The residue was taken up in methanol, and the precipitate was filtrated off, rinsed with methanol and dried to give RNA_{*x+1*}-ACSS as a white solid.

5'-*O*-Dimethoxytrityl-*N*³,2'-*O*-dimethyluridine (5)

Compound **5** was prepared in two steps in 71% overall yield from commercially available 2'-*O*-methyluridine by *N*-

3 methylation with methyl iodide (29) followed by 5'-*O*-dimethoxytritylation. ¹H and ¹³C NMR spectra are in accordance with the literature (30).

5'-*O*-Dimethoxytrityl-3'-*O*-succinyl-*N*³,2'-*O*-dimethyluridine (6)

A mixture of compound **5** (0.50 g, 0.87 mmol), anhydrous succinic anhydride (0.13 g, 1.31 mmol), 4-dimethylaminopyridine (0.05 g, 0.44 mmol) and triethylamine (0.12 ml, 0.87 mmol) in dichloromethane (5 ml) was stirred for 2 h at room temperature. After addition of dichloromethane, the organic layer was washed with 2% citric acid, water and then dried with anhydrous sodium sulfate and concentrated under vacuum to give compound **6** (0.55 g, 88%). ¹H NMR (400 MHz, DMSO-*d*₆): δ 2.46–2.60 (m, 4H, CH₂), 3.17 (s, 3H, CH₃N), 3.26–3.40 (m, 5H, CH₃O, H-5', H-5''), 3.75 (s, 6H, CH₃O), 4.13–4.18 (m, 1H, H-4'), 4.20 (t, *J* = 4.9 Hz, 1H, H-2'), 5.26 (t, *J* = 5.6 Hz, 1H, H-3'), 5.56 (d, *J* = 8.0 Hz, 1H, H-5), 5.86 (d, *J* = 4.4 Hz, 1H, H-1'), 6.87–6.96 (m, 4H, H-arom), 7.20–7.28 (m, 5H, H-arom), 7.29–7.40 (m, 4H, H-arom), 7.78 (d, *J* = 8.0 Hz, 1H, H-6 pyrimidine); ¹³C NMR (100 MHz, DMSO-*d*₆): δ 27.7, 29.2, 39.4, 39.6, 39.8, 40.0, 40.3, 40.5, 40.7, 55.3, 55.5, 58.7, 62.8, 70.5, 74.0, 80.7, 80.9, 86.6, 88.8, 101.5, 113.8, 127.3, 128.1, 128.4, 130.2, 135.5, 135.7, 139.2, 145.0, 151.1, 158.7, 162.4, 172.0, 173.7; HRMS (ESI-TOF) *m/z*: calcd for [C₃₆H₃₈N₂O₁₁+Na]⁺ 697.2373, found: 697.2395.

3'-*O*-[4-oxo-4-(4-(3,4,5-tris(octadecyloxy)benzoyl)piperazin-1-yl)]-succinyl-*N*³,2'-*O*-dimethyluridine (7, namely RNA₁-ACSS)

To compound **6** (0.74 g, 0.72 mmol) in dichloromethane (5 ml) were added ACSS **4** (0.50 g, 0.74 mmol), *O*-benzotriazole-*N,N,N'*-tetramethyl-uronium-hexafluoro-phosphate (HBTU) (0.67 g, 2.39 mmol), 1-hydroxybenzotriazole (HOBT) (0.32 g, 2.13 mmol) and *N,N*-diisopropylethylamine (DIEA) (0.38 ml, 2.16 mmol). After stirring for 1 h at room temperature, the reaction mixture was diluted with methanol (10 ml), concentrated and the precipitate was filtrated off, rinsed with methanol and dried under vacuum. The precipitate was taken up in dichloromethane (40 ml) and a solution of 5% DCA in dichloromethane (12 ml) was added. After stirring for 1 h, the reaction mixture was diluted with methanol (25 ml) and concentrated under vacuum. The precipitate was filtrated off, rinsed with methanol and dried to give RNA₁-ACSS (**7**) as a white solid (0.61 g, 0.45 mmol, 63%). ¹H NMR (400 MHz, CDCl₃): δ 0.90 (t, *J* = 6.7 Hz, 9H, OCH₂CH₂CH₂(CH₂)₁₄CH₃), 1.19–1.41 (m, 84H, OCH₂CH₂CH₂(CH₂)₁₄CH₃), 1.42–1.55 (m, 6H, OCH₂CH₂CH₂(CH₂)₁₄CH₃), 1.72–1.92 (m, 6H, OCH₂CH₂CH₂(CH₂)₁₄CH₃), 2.6–2.85 (m, 4H, CH₂ succinyl), 3.35 (s, 3H, NCH₃), 3.48 (s, 3H, OCH₃), 3.51–3.77 (m, 8H, CH₂ piperazine), 3.78–3.87 (m, 1H, H-5'), 3.92–4.04 (m, 7H, H-5'' and OCH₂CH₂CH₂(CH₂)₁₄CH₃), 4.22–4.29 (m, 2H, H-2' and H-4'), 5.31 (t, *J* = 4.7 Hz, 1H, H-3'), 5.78 (d, *J* = 4.3 Hz, 1H, H-1'), 5.81 (d, *J* = 8.2 Hz, 1H, H-5), 6.61 (s, 2H, H arom), 7.71 (d, *J* = 8.2 Hz, 1H, H-6); ¹³C NMR (100 MHz, CDCl₃): δ 14.0, 22.6, 26.1,

27.6, 28.0, 29.1, 29.3, 29.3, 29.4, 29.5, 29.6, 29.7, 30.3, 31.9, 58.9, 61.4, 69.4, 70.5, 73.5, 76.7, 77.0, 77.2, 77.3, 81.2, 82.9, 90.7, 102.0, 105.9, 129.7, 139.0, 139.7, 151.1, 153.3, 162.6, 169.7, 170.7, 172.7; Maldi-TOF-MS: m/z calcd for $[C_{80}H_{140}N_4O_{12}+Na]^+$ 1372.036, found: 1372.079.

RNA₂-ACSS. Coupling of **RNA₁-ACSS (7)** (0.54 g, 0.40 mmol) and 5'-*O*-dimethoxytrityl-2'-*O*-methyl-*N*-acetyl-cytidine-3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (**8**) (0.80 g, 2.5 eq) following the general procedures afforded **RNA₂-ACSS** (0.64 g, 0.36 mmol, 94%). Maldi-TOF-MS: m/z calcd for $[C_{95}H_{159}N_8O_{20}P+Na]^+$ 1786.131, found 1786.142.

RNA₃-ACSS. Coupling of **RNA₂-ACSS** (0.64 g, 0.36 mmol) and 5'-*O*-dimethoxytrityl-2'-*O*-methyl-*N*-benzoyl-adenosine-3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (**9**) (0.80 g, 2.5 eq) following the general procedures afforded **RNA₃-ACSS** (0.73 g, 0.32 mmol, 89%). Maldi-TOF-MS: m/z calcd for $[C_{116}H_{180}N_{14}O_{27}P_2+Na]^+$ 2286.251, found 2286.263.

RNA₄-ACSS. Coupling of **RNA₃-ACSS** (0.72 g, 0.32 mmol) and 5'-*O*-dimethoxytrityl-6-*N*,6-*N*,2'-*O*-trimethyl-adenosine-3'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (**10**) (see Supplementary Information) (0.72 g, 2.8 eq) following the general procedures afforded **RNA₄-ACSS** (0.74 g, 0.27 mmol, 84%). Maldi-TOF-MS: m/z calcd for $[C_{132}H_{201}N_{20}O_{33}P_3+Na]^+$ 2710.377, found: 2710.402.

DMT-P-RNA₄-ACSS. Coupling of **RNA₄-ACSS** (0.31 g, 0.115 mmol) and 2-[2-(4,4'-dimethoxytrityloxy)ethylsulfonyl]ethyl-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramidite (3.5 eq) following the general procedure A afforded **DMT-P-RNA₄-ACSS** (0.37 g, 0.114 mmol, 99%). Maldi-TOF-MS: m/z calcd for $[C_{160}H_{231}N_{21}O_{41}P_4S+Na]^+$ 2979.390, found: 2979.701.

5'-Phosphorylated tetramer (2)

Cleavage of RNA from the support and base/phosphate deprotection were performed by treatment of **DMT-P-RNA₄-ACSS** (42.6 mg, 0.015 mmol) with 28% aqueous ammonia in ethanol (1.0 ml, 3:1 v/v) for 90 min at 80°C. After lyophilization, the residue was taken up twice in water, centrifuged and the supernatant was filtered and purified by reverse-phase HPLC (5–20% linear gradient of ACN in 10 mM TEAA buffer over 20 min) to give **P-RNA₄** (**2**) as 2.8 equiv triethylammonium salt (12.4 mg, 0.009 mmol, 61%). ¹H NMR (400 MHz, D₂O): δ 1.26 (t, J = 7.5 Hz, 24H, CH₃ NEt₃), 3.13–3.36 (m, 26H, OCH₃ and CH₂ NEt₃), 3.49 (s, 3H, OCH₃), 3.54 (s, 3H, OCH₃), 3.66 (s, 3H, OCH₃), 3.70 (s, 3H, OCH₃), 3.74–4.92 (m, 20H, H-2', H-3', H-4', H-5' and H-5''), 5.68 (br s, 1H, H-1'), 5.70 (d, J = 7.7 Hz, 1H, H-5 rU), 5.84 (d, J = 8.1 Hz, 1H, H-5 rC), 5.95–6.00 (m, 2H, H-1'), 6.11 (d, J = 2.3 Hz, 1H, H-1'), 7.69 (d, J = 7.7 Hz, 1H, H-6 rU), 7.83 (d, J = 8.1 Hz, 1H, H-6 rC), 7.89 (s, 1H, H-2 purine), 8.13 (s, 2H, H-8 purine), 8.38 (s, 1H, H-2 purine);

³¹P NMR (169 MHz, D₂O): δ -0.23, -0.10, 0.03, 1.51; Maldi-TOF-MS: m/z calcd for $[C_{45}H_{63}N_{15}O_{28}P_4+H]^+$ 1386.299, found: 1386.339.

P²-Imidazolidine-*N*⁷-methylguanosine 5'-*O*-diphosphate (3)

A mixture of triethylammonium salt of 7-methylguanosine 5'-diphosphate (157 mg, 0.115 mmol) [synthesized from GDP by reaction of methyl iodide in DMSO according to reported procedures (29), imidazole (155 mg, 2.28 mmol), 2,2'-dithiodipyridine (128 mg, 0.90 mmol), triethylamine (157 μ l) and triphenylphosphine (240 mg, 0.92 mmol) in dimethylformamide (1.6 ml) was stirred overnight at room temperature. A solution of sodium perchlorate (100 mg, 0.80 mmol) in anhydrous acetone (4 ml) was then added at 4°C. The precipitate recovered by centrifugation was washed with four portions of cold acetone, and dried to give Im(m⁷GDP) (**9**) as sodium salt (50.9 mg, 92 μ mol, 80%). NMR spectra are in accordance to literature data (30). HRMS (ESI-TOF): m/z calcd for $[C_{13}H_{20}N_5O_7P-H]^-$ 388.1022, found: 388.1016.

Trypa cap-4 (1)

A mixture of the triethylammonium salt of **2** (12.9 mg, 7.74 μ mol), the sodium salt of Im(m⁷GDP) **3** (27.0 mg, 49 μ mol) and anhydrous ZnCl₂ (14 mg, 105 μ mol) in dimethylformamide (1.2 ml) was stirred overnight at room temperature. The reaction mixture was then poured into 1.8 M aqueous solution of EDTA (10 ml) and pH was adjusted to 6.5 with 1M Na₂CO₃. Preparative HPLC purification (10–17% linear gradient of acetonitrile in 10 mM TEAA over 14 min) gave trypa cap-4 (**1**) as triethylammonium salt (5.9 mg, 2.95 μ mol, 40%). To tetramer **1** dissolved in water (80 μ l) was added slowly at 4°C a cooled solution of sodium perchlorate (40 mg) in anhydrous acetone (800 μ l). The precipitate was recovered by centrifugation, washed with four portions of cold acetone and then dried (5.15 mg, 2.66 μ mol, 34%). ³¹P NMR (169 MHz, D₂O): δ -0.23, -0.10, 0.03, 1.49; HRMS (ESI-TOF): m/z calcd for $[C_{56}H_{79}N_{20}O_{38}P_6-3H]^{3-}$ 607.0992, found: 607.1012.

Protein Expression and Purification

The nucleotide sequences encoding the full-length and a truncated construct comprising residues 1 to 189 of TcEIF4E5 (EKG01513.1) were synthesized by Biomatik, after codon optimization for *Escherichia coli* expression system, and cloned into the pET28a vector (Novagen). In this vector, the recombinant proteins are expressed in fusion with a N-terminal His-tag. *E. coli* BL21 Star (DE3) cells were transformed with the expression vector and incubated at 37°C in LB medium containing the selection antibiotics. When the cultures reached OD₆₀₀ ~0.6, the temperature was reduced to 18°C and protein expression was induced with 0.25 mM of isopropyl- β -D-thiogalactopyranoside (IPTG) for ~16 h. Cells from 1 l culture were harvested by centrifugation at 6000 \times g for 15 min, suspended and lysed in 20 ml of buffer A (50 mM Tris-HCl pH 8, 200 mM NaCl, 20 mM imidazole, 1 mM de DTT) using a microfluidizer (Microfluidics™ 110L series 0300).

The soluble fractions were isolated by centrifugation at $40\,000 \times g$ for 30 minutes at 4°C. The extract was loaded onto a 1 ml His-Trap HP column (GE Healthcare Life Sciences) equilibrated with buffer A. TcEIF4E5 was eluted with a 10 column volumes (CV) linear gradient from 0 to 15% buffer B (buffer A + 500 mM imidazole), followed by a 15 CV linear gradient from 15 to 100% buffer B. TcEIF4E5-ΔC was eluted with a 10 column volumes (CV) linear gradient from 0 to 10% buffer B, followed by a 15 CV linear gradient from 10 to 100% buffer B. Fractions containing the target protein were pooled, concentrated to a final volume of 2 ml and loaded onto a Superdex 75 16/60 (GE Healthcare Life Sciences) equilibrated with 20 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM DTT. TcEIF4E5 and TcEIF4E5-ΔC purified by this procedure were used in microscale thermophoresis assays.

Purification of TcEIF4E5 for crystallization assays followed a slightly different protocol. pDEST17 expression vector (Invitrogen) containing the gene encoding TcEIF4E5 sequence was kindly provided by the laboratory of Dr. Fabiola Holetz (Instituto Carlos Chagas, FIOCRUZ, Brazil) and used to transform *E. coli* BL21 (DE3) ΔSlyD_pRARE2. Protein expression and affinity chromatography followed the same procedures described above. After Ni²⁺ affinity chromatography, the fractions containing the target protein were pooled, dialyzed against 50 mM Tris-HCl pH 8, 20 mM NaCl, 2 mM DTT and loaded onto a 5 ml Hi-Trap SP HP column (GE Healthcare Life Sciences). TcEIF4E5 was eluted with a 20 CV linear gradient from 20 mM to 1 M NaCl. The sample was concentrated to a final volume of 2 ml and loaded onto a Superdex 75 16/60 (GE Healthcare Life Sciences) equilibrated with 20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM DTT. The eluted protein was concentrated to ~5 mg/ml for crystallization assays.

Expression of selenomethionine-labeled (SeMet) TcEIF4E5 was performed based on the metabolic inhibition method (31). Briefly, *E. coli* BL21 Star (DE3) cells transformed with the expression vector were incubated at 37°C, 200 rpm in M9 minimal medium supplemented with the selection antibiotic. When the culture reached OD₆₀₀ ~0.6, it was supplemented with 10 ml of a sterile mixture containing lysine, phenylalanine and threonine (10 mg/ml), isoleucine, leucine and valine (5 mg/ml) and selenomethionine (6 mg/ml) diluted into M9 medium without glucose and thiamine. The temperature was reduced to 18°C and after 30 min of incubation, expression was induced with 0.25 mM IPTG for approximately 16 h. Cells were harvested by centrifugation and TcEIF4E5-SeMet was purified using the same protocol applied to non-labelled TcEIF4E5.

Crystallization, data collection and processing

TcEIF4E5 was incubated with m⁷GTP (Jena Bioscience), cap-1 and cap-4 at a molar ratio of 1:1.5 (protein:ligand). Protein samples in the presence and absence of ligands were submitted to crystallization trials by sitting drop vapor-diffusion method using commercial screens. Initial hits were obtained using TcEIF4E5 after incubation with cap ana-

logues but not in the absence of ligands. Optimization of the initial crystallization conditions was performed by varying precipitant and protein concentrations. Best crystals were obtained by hanging drop vapor diffusion at 18°C by mixing the TcEIF4E5 protein at 5 mg/ml (0.2 mM) in 20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM DTT with crystallization buffer containing 100 mM BIS-TRIS pH 6 or 6.5, 200 mM Lithium sulfate, 25–30% PEG 3350. TcEIF4E5-SeMet at 5 mg/ml in 20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM DTT was crystallized in buffer containing 20 mM ammonium sulfate, 10 mM HEPES pH 7.5, 25% PEG 3350.

Before data collection, the crystals were cryo-protected with 20% (v/v) glycerol added to the mother liquor prior to flash cooling in liquid nitrogen. X-ray diffraction data from TcEIF4E5-m⁷GTP crystals were collected at the PROXIMA 1 beam line of the Synchrotron SOLEIL using a PILATUS 6M detector (Dectris). Diffraction data from TcEIF4E5-cap1 and TcEIF4E5-cap4 crystals were collected at the same beam line using an EIGER X 16M detector (Dectris). X-ray diffraction data from TcEIF4E5-SeMet crystals were collected at the PROXIMA 2 beam line of the Synchrotron SOLEIL using an EIGER X 9M detector (Dectris) at the Se-absorption wedge wavelength (0.97830 Å). Diffraction data were processed with XDS package (32). Data statistics are presented in Table 1 and Supplementary Table S1.

Structure determination and refinement

The structure of TcEIF4E5 was determined by single-wavelength anomalous diffraction using diffraction data from TcEIF4E5-SeMet. The heavy atoms' positions were determined using the program SHELXD (33), the initial phases were calculated using PHASER (34) and density modification was performed using PARROT (35). The low quality of the experimental electron density map hindered the construction of an initial model, either manually with COOT (36) or automatically using the program BUCCANEER (37). Therefore, we used MOLREP (38) to fit a homology model generated using the Phyre2 server (39) into the electron density map. After model edition using COOT to keep only selected fragments which were in good accordance with the experimental electron density map, we used this partial model together with the heavy atoms' positions to recalculate phases using PHASER, followed by density modification with PARROT. The new electron density map allowed improvement of the model construction using COOT. At this point, we used the higher resolution data set of the TcEIF4E5-m⁷GTP crystal in order to proceed with model building and structure refinement. TcEIF4E5-m⁷GTP refined structure was used as model for TcEIF4E5-cap1 and TcEIF4E5-cap4 structure determination by molecular replacement. Refinement of the structures were performed alternating cycles of BUSTER (40) with visual inspection and manual rebuilding using COOT. The geometrical restraints for cap-1 and cap-4 molecules were generated using the Grade Web Server (<http://grade.globalphasing.org>) (41). Model validation was performed using the MolProbity Web Server (<http://molprobity.biochem.duke.edu/>) (42).

Table 1. Crystallographic data and refinement statistics

Crystal	TcIF4E5-m ⁷ GTP	TcIF4E5-cap1	TcIF4E5-cap4
<i>Data statistics</i>			
Source	SOLEIL-PX1	SOLEIL-PX1	SOLEIL-PX1
Space group	<i>P</i> 3 ₁ 21	<i>P</i> 3 ₁ 21	<i>P</i> 3 ₁ 21
Unit cell (Å)	<i>a</i> = <i>b</i> = 100.72, <i>c</i> = 40.63	<i>a</i> = <i>b</i> = 102.82 <i>c</i> = 42.50	<i>a</i> = <i>b</i> = 102.52, <i>c</i> = 41.11
Resolution (Å)	50–2.10 (2.23–2.10)	50–2.70 (2.86–2.70)	50–2.20 (2.33–2.20)
Number of observations	204 339 (32 204)	149 068 (23 611)	262 446 (41 327)
Number of unique reflections	13 943 (2205)	7339 (1176)	12 900 (2040)
Completeness (%)	99.2 (98.2)	100.0 (99.9)	100.0 (99.9)
Redundancy	14.6 (14.6)	20.3 (20.1)	20.3 (20.3)
R _{meas} (%)	6.9 (262.6)	6.6 (262.6)	6.4 (204.7)
I/σ	19.9 (1.05)	24.1 (0.98)	23.7 (1.01)
CC $\frac{1}{2}$	100.0 (58.1)	100.0 (54.4)	100.0 (58.7)
<i>Refinement statistics</i>			
R _{work} /R _{free} (%)	17.8/19.8	19.4/25.2	20.7/22.6
RMSD bond length(Å)/angle (°)	0.01/1.09	0.01/1.19	0.01/1.15
Mean B value (overall, Å ²)	84.7	136.6	93.3
Protein atoms	1414	1447	1452
Heterogen atoms	38	59	130
Number of water molecules	72	23	68
<i>Ramachandran plot</i>			
Favored (%)	98.3	96.1	98.3
Outliers (%)	0	0	0

Values in parentheses are for the highest resolution shell.

Microscale thermophoresis

TcEIF4E5 and TcEIF4E5-ΔC were labeled using the Monolith NT Protein Labeling Kit RED-Tris-NTA (NanoTemper Technologies) according to the supplier instructions. The concentration of labeled proteins was kept constant at 12.5 nM. The unlabeled binding partners (m⁷GTP or cap-4) were titrated in 1:1 dilutions with the highest final concentration of 2.5 mM. The assays were performed in triplicate in buffer containing 20 mM Tris-HCl pH 8, 100 mM NaCl, 2 mM DTT and 0.05% Tween-20. Measurements were performed in treated capillaries (NanoTemper Technologies) on a Monolith NT.115 Pico system (Nano Temper Technologies) using 20% LED and Medium IR-laser power. Data analysis were performed using MO.affinity analysis v.2.2.7 software (NanoTemper Technologies).

RESULTS

Large scale synthesis of cap-4

An efficient and robust synthesis protocol for synthetic route to trypanosome cap-4 (m⁷Gpppm₃^{6,6,2} Apm² Apm² Cpm₂^{3,2} U) (compound **1**, Figure 1) was developed by taking advantage of recent developments in liquid-phase oligonucleotide synthesis (LPOS) which combine the beneficial aspects of both conventional solid-phase and solution-phase chemistry (43). We have also produced the cap-1 analogue (m⁷Gpppm₃^{6,6,2} A) for comparative structural studies.

As illustrated in Figure 1, the key step in trypa cap-4 synthesis is the formation of the triphosphate linkage by coupling the phosphorylated tetramer P-RNA₄ (**2**) with the m⁷GDP imidazolidine (**3**) in the presence of ZnCl₂ as catalyst (44). The 5'-phosphorylated tetramer (**2**, P-RNA₄) was prepared by liquid-phase synthesis on an alkyl-chain-

soluble support (ACSS, **4**) using phosphoramidite chemistry. The soluble support **4** was obtained in 5 steps and 56% overall yield from commercially available methyl gallate (43). Two of the four RNA building blocks (**5** and **10**, Figure 2) are not commercially available. The detailed procedures for their synthesis and analytical data can be found in the Supplementary Data. The first modified nucleoside, 5'-*O*-dimethoxytrityl-*N*³,2'-*O*-dimethyluridine (**5**), was obtained in two steps from 2'-*O*-methyluridine by *N*-3 methylation with methyl iodide (29), followed by 5'-*O*-dimethoxytritylation. The fourth nucleoside phosphoramidite **10** was obtained in six steps from 6-chloropurine riboside using slight changes in established procedures (30). Briefly, amination of 6-chloropurine riboside with dimethylamine was achieved by microwave irradiation at 95°C for 30 min. 3'-*O*-Methylation was performed with methyl iodide in DMSO after selective protection of the 3' and 5'-hydroxyl groups using the Markiewicz's reagent. Deprotection of silyl ethers, followed by 5'-*O*-dimethoxytritylation and 3'-*O*-phosphitylation according to classical protocols, afforded the phosphoramidite **10** in an overall yield of 25% in six steps (Supplementary Scheme S1).

RNA synthesis started with attachment of the nucleoside **5** to the soluble support via a succinate linkage. The resulting anchored nucleoside **7** was then submitted to phosphoramidite coupling cycle using the building blocks **8**, **9** and **10** under optimized conditions (43). The last step of the synthesis consisted of the 5'-phosphorylation of the tetramer still anchored to the support by using the commercially available phosphitylating reagent (2-[2-(4,4'-dimethoxytrityloxy)ethylsulfonyl]ethyl-(2-cyanoethyl)-(N,N-diisopropyl)phosphoramidite). Removal of all protecting groups and cleavage from the support were achieved by conventional ammonolysis. The fully deprotected tetramer 5'-monophosphate (**1**) was isolated after reverse phase HPLC purification in high yield and

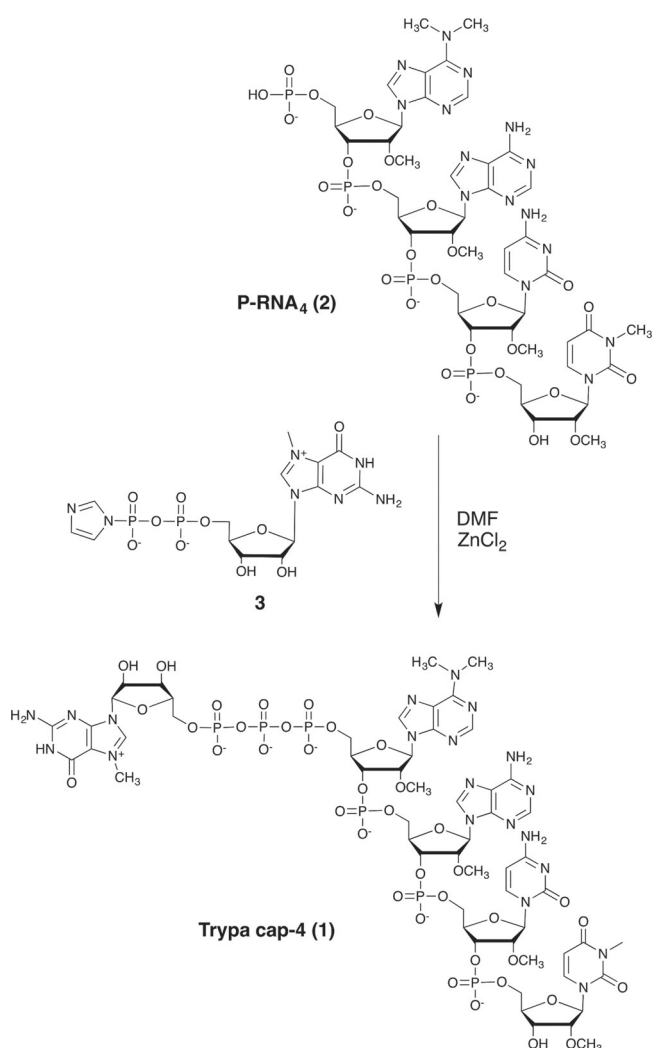


Figure 1. Capping step of 5'-phosphorylated tetramer **2** with imidazolidine **3** in the presence of ZnCl₂ as catalyst.

purity. Finally, the phosphorylated tetramer **2** was coupled to 7-methylguanosine 5'-diphosphate imidazolidine (**3**) in the presence of anhydrous zinc chloride as catalyst (30,44). Trypa cap-4 (**1**) was isolated as sodium salt in 34% yield. The chemical structure and purity of **1** were confirmed by mass spectrometry and HPLC analysis. Cap-1 (m⁷Gpppm₃^{6,6,2}, **17**) was obtained by coupling N⁶,N⁶,2'-O-trimethyladenosine 5'-monophosphate (**16**) and phosphoimidazolidine **3** (Supplementary Scheme S2).

The crystal structure of *T. cruzi* EIF4E5 in complex with cap analogs and cap-4 binding mechanism

Initial attempts to determine TcEIF4E5 structure by molecular replacement methods failed, although several homologue structures are available in the Protein Data Bank. Therefore, selenomethionine-labelled (SeMet) TcEIF4E5 was produced and experimental phases were obtained by single wavelength anomalous diffraction. As TcEIF4E5-SeMet crystals showed poor diffraction, the partial model

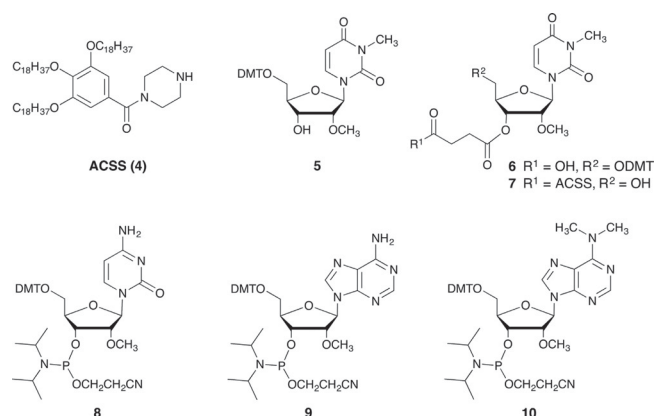


Figure 2. Building blocks **4–10** used for RNA assembly.

obtained from experimental electron density maps was used as an initial model to complete building and refinement against a higher resolution native data set (see Materials and Methods).

The TcEIF4E5 structure in complex with 7-methyl-GTP (m⁷GTP) was refined at 2.10 Å resolution to final $R_{\text{factor}}/R_{\text{free}}$ of 17.8%/19.8%. The TcEIF4E5 structures in complex with cap-1 and cap-4 were refined at 2.70 and 2.20 Å resolution to final $R_{\text{factor}}/R_{\text{free}}$ of 19.4%/25.2% and $R_{\text{factor}}/R_{\text{free}}$ of 20.7%/22.6%, respectively (Table 1). The models comprise residues 4–186 with the polypeptide chain clearly defined by the electron density. The exception is the region connecting β 1– β 2 strands, which participates in the cap binding and is better defined in TcEIF4E5-cap-4 structure. Residues 30 to 32 could not be modeled in TcEIF4E5-m⁷GTP complex whereas residues 30–31 are lacking in TcEIF4E5-cap-1 model (Figure 3). Superposition of TcEIF4E5 structures resulted in RMSD values of 0.37 Å. TcEIF4E5 conserves the overall structure of the eIF4E family, a curved-shape α/β protein with the cap binding site located on the concave surface of the protein (Figure 3A). The cap analogues were well defined by the electron density maps, except for the base moiety of cap-1 second nucleotide (Figure 3B–D). The second ribose and base moieties of cap-1 adopt a distinct conformation compared with cap-4, probably due to flexibility of the cap-1 analogue in the TcEIF4E5-cap-1 complex (discussion below).

The structure of the TcEIF4E5-cap-4 complex revealed an extensive protein-ligand interaction surface. Binding of cap-4 is mediated by hydrophobic interactions as well as hydrogen bonds (Figure 4). TcEIF4E5 conserves the classical stacking interaction with the m⁷G moiety, with W33 and Y83 (at equivalent positions of human eIF4E1 W56 and W102) forming the m⁷G base sandwich. The highly conserved adjacent glutamic acid (E84) forms two hydrogen bonds with N1 and N2 of the guanine base and Y83 main chain-amide forms a H-bond with the guanine C-6 carbonyl group. The arginine residues R91, R133, R137 are involved in hydrogen bonds with the phosphate groups of m⁷GTP. In addition, the m⁷GTP ribose moiety is involved in hydrophobic interactions with L20, P22, V25 and main chain of T21, located in the region connecting β 1– β 2 strands. Interestingly, the cap-4 AACU bases pack together in a par-

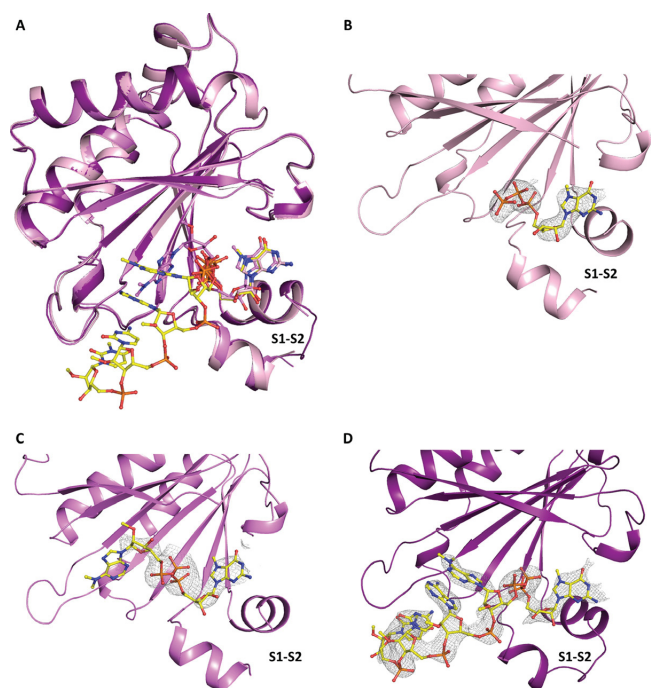


Figure 3. Three-dimensional structure of *T. cruzi* EIF4E5. (A) Overall superposition of TcEIF4E5 in complex with m⁷GTP, cap-1 and cap-4 (light pink, magenta and purple, respectively). (B–D) Electron density maps corresponding to the cap analogues. The ligands m⁷GTP (B), cap-1 (C) and cap-4 (D) are clearly defined by the electron density, except for the base moiety of cap-1 second nucleotide. $2F_o - F_c$ electron density maps are represented in gray and contoured at 1.0σ . The ligands are represented in sticks. The region connecting β1–β2 strands (S1–S2) is indicated.

allel stacking conformation (Figure 4B), oriented towards the protein, leaving the phosphate groups exposed to the solvent. Besides stacking between the bases, the first methylated adenine is involved in additional hydrophobic interactions with the aliphatic moieties of residues D95, L145 and R147. The second adenine base is involved in hydrophobic interactions with N139, in addition to two hydrogen bonds with R137 main chain-carbonyl and N139 side chain. P138 main chain-carbonyl and Q140 main chain-amide form H-bonds with the cytosine base, whereas the last nucleotide (U) is only involved in hydrophobic interactions. Figure 4C shows cap-4–TcEIF4E5 interactions in detail. Methylation of cap-4 bases seems to be determinant of its structural arrangement in the TcEIF4E5–cap-4 complex. Notably, the N6 and N3 methyl groups of adenine and uracil are involved in hydrophobic interactions with TcEIF4E5 residues, besides having a role in stacking of AACU bases. The 2'-O-methyl groups of cytidine and uridine also participate in hydrophobic interactions involving the TcEIF4E5 loop connecting β5–β6 strands, whereas the first two adenosine ribose methyl groups are exposed to the solvent.

Unique molecular features mediate *T. cruzi* EIF4E5–cap-4 interaction

Superposition of the TcEIF4E5 structure (cap-4 complex) with mammalian eIF4E family members resulted in RMSD varying from 1.90 to 2.77 Å (146 and 157 C-alpha atoms

aligned with human eIF4E1 and eIF4E2 (PDB: 1IPC and PDB: 2JGB) respectively, 157 C-alpha atoms aligned with mouse eIF4E3 (PDB: 4B6U)). The main structural differences of TcEIF4E5 compared with eIF4E family members reside in the regions connecting β1–β2, β3–β4 and β5–β6 strands (Figure 5A). The first one, also called S1–S2 loop, comprises one of the residues that forms the m⁷G sandwich (W33 in TcEIF4E5, W56 in human eIF4E1) and is described to undergo structural rearrangement upon cap binding (14). The corresponding region in TcEIF4E5 is longer, comprising 19 residues, in contrast to a 7-residues loop present in human eIF4E1. In TcEIF4E5, two short α-helices are formed, neighboring residue W33. Human eIF4E2 and mouse eIF4E3 also present one helix in S1–S2 (14,15), however their β1–β2 connections are significantly shorter compared with TcEIF4E5 (Figure 5). The second residue forming the m⁷G sandwich (Y83 in TcEIF4E5, W102 in human eIF4E1) is located in the region connecting β3–β4 strands (S3–S4 loop), which also shows a conformational variation among the structures. Finally, the region connecting β5–β6 strands is longer in TcEIF4E5, encompassing 9 residues, in contrast to a 4-residues loop in the other counterparts (Figure 5A). Interestingly, the TcEIF4E5 extended S5–S6 loop creates an additional molecular surface that mediates the interaction with the cap-4 AACU bases (Figure 4A).

A detailed comparison of the residues involved in TcEIF4E5 cap recognition reveals further differences from other eIF4E family members. Although substitution of the tryptophan residue located at the flexible S1–S2 region for other aromatic or non-aromatic residue has been previously described, this is the first eIF4E homologue structure showing a substitution of the tryptophan equivalent to human W102 located at the S3–S4 loop (TcEIF4E5 Y83) (Figure 5B–D). In addition, an adjacent methionine (human M101) also involved in hydrophobic interactions with the guanine base is replaced by a serine (S82) in TcEIF4E5. The correct positioning of less bulky residues Y83 and S82 to interact with the m⁷G moiety in TcEIF4E5 was most likely made possible by the flexibility of S3–S4 loop.

Interestingly, TcEIF4E5 W33, located at the beginning of the S1–S2 second short helix, interacts with the cap-4 guanine base in a distinct orientation, with the pyrrole ring stacking with the pyrimidine moiety of the m⁷G (Figure 5B). The extended β1–β2 connection observed in TcEIF4E5 creates additional surface for cap interaction with an additional short helix, placing residues V25, P22, L20 in position to participate in hydrophobic interactions with the first ribose moiety. The implication of aliphatic and backbone residues of the S1–S2 loop to form the cap pocket is also observed in mouse eIF4E3 (Figure 5D). However, in the case of mouse eIF4E3 this additional hydrophobic surface is shorter and proposed to compensate the lack of the S1–S2 aromatic residue to form the m⁷G sandwich (15). Similar to human eIF4E1 and eIF4E2, basic side chains are involved in TcEIF4E5 interactions with the m⁷GTP phosphate groups (Figure 5B, C). However, in striking contrast with other eukaryotic counterparts, very few water molecules are present in the cap-4 pocket (Supplementary Figure S1). In the TcEIF4E5–cap-4 complex there is only three water molecules participating in the cap-binding inter-

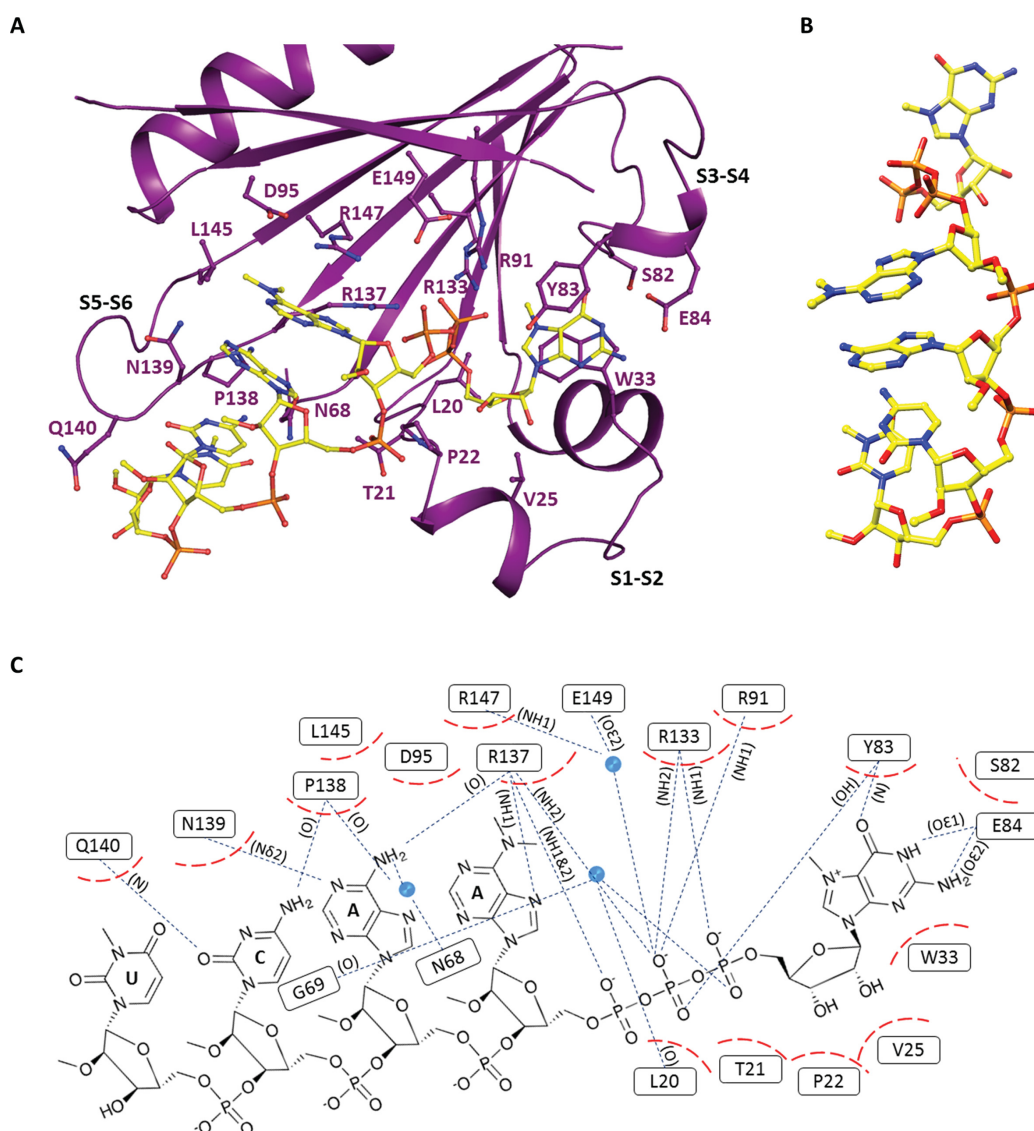


Figure 4. *T. cruzi* EIF4E5 cap recognition site. (A) Representation of the cap-4 binding site. Residues involved in the interaction with cap-4 are shown in sticks and labeled. The regions connecting $\beta 1$ – $\beta 2$, $\beta 3$ – $\beta 4$ and $\beta 5$ – $\beta 6$ strands (S1–S2, S3–S4 and S5–S6) are indicated. (B) Three-dimensional structure of cap-4 evidencing the AACU base stacking. (C) Schematic representation showing the cap-4-TcEIF4E5 interactions in detail. Hydrophobic interactions are represented in dashed red and hydrogen bonds in blue dotted lines. The atoms involved in the hydrogen bonds are indicated in parentheses. Water molecules are represented as blue spheres.

face, two of them mediating phosphate interactions with the protein side chains (Figure 4B). In the TcEIF4E5- m^7 GTP complex, seven water molecules are involved in the hydration of the cap pocket, four of them participating in phosphates interaction (see Discussion).

Comparison of *T. cruzi* EIF4E5 and *Leishmania* EIF4E1 structures provides insights into cap-4 binding specificities of trypanosomatids EIF4E homologues

Superposition of the TcEIF4E5 and *L. major* EIF4E1 (PDB 5wb5) structures resulted in an RMSD of 2.06 Å for 157 C- α atoms aligned. Significant differences were observed in the cap-binding pocket of these two trypanosomatid ho-

mologues (Figure 6). The LmEIF4E1 structure was determined in its apo form showing tryptophans W37 and W83, that are expected to sandwich the cap m^7 G-moiety in a flipped conformation (28). The LmEIF4E1 S1–S2 loop is shorter than in TcEIF4E5 (9 residues instead of 19), but a structural rearrangement upon cap-binding could bring a proline containing segment (28-SPST-31) in position to form a hydrophobic interacting surface with the cap ribose moiety, similar to that observed in TcEIF4E5 (Figure 6B). Basic side chains in position to interact with the m^7 GTP phosphates (K93, R167, R172) are also present in LmEIF4E1 (Figure 6B). However, the S5–S6 loop that forms the surface of interaction with the cap-4 AACU bases

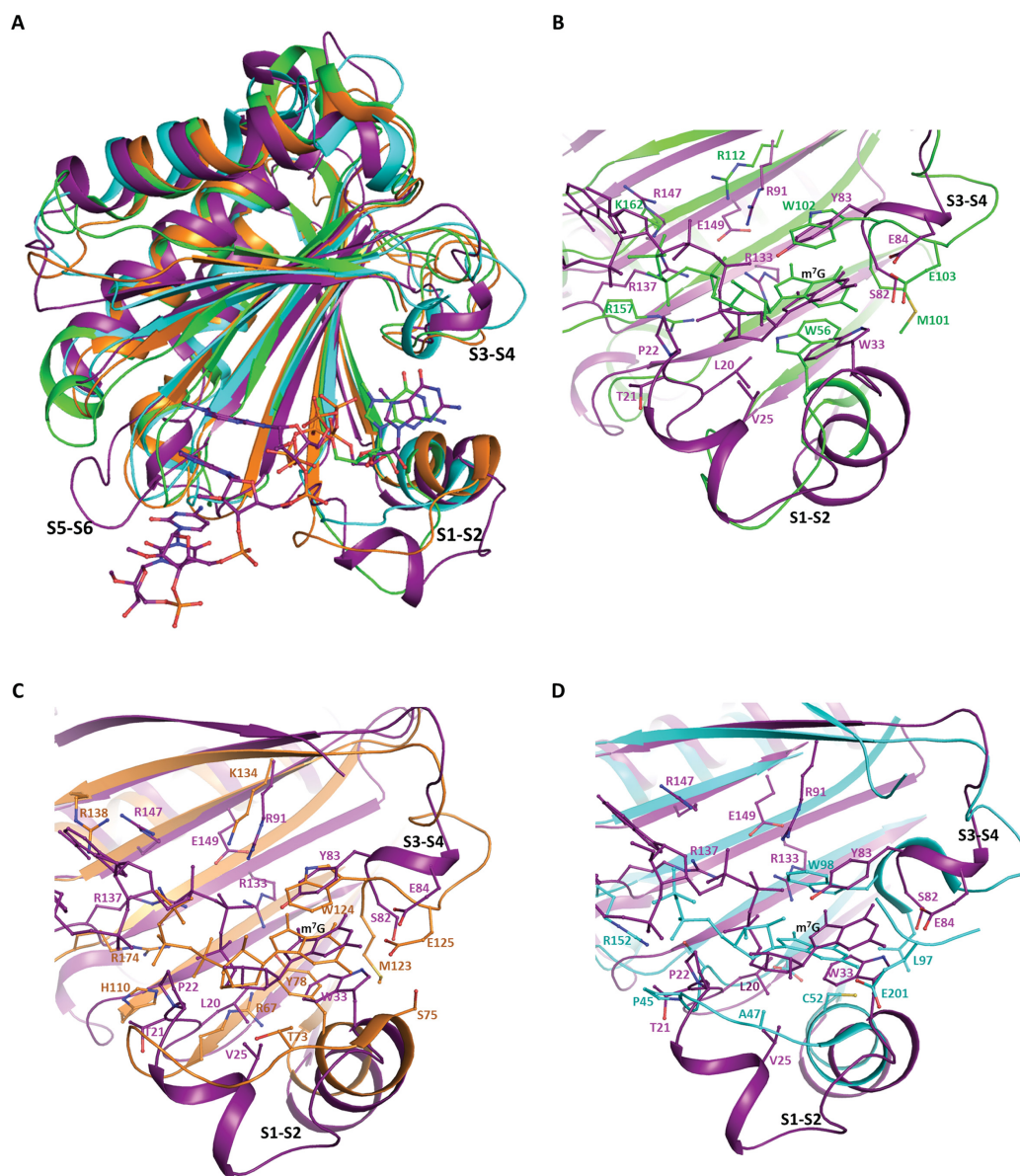


Figure 5. Structural comparison of *T. cruzi* EIF4E5 with other eukaryotic IF4E factors. (A) Overall superposition: TcEIF4E5 is shown in purple, human eIF4E1 and eIF4E2 (HsIF4E1, PDB 1IPC; HsIF4E2, PDB 2JGB) are represented in green and orange, respectively. Mouse eIF4E3 (MmIF4E3, PDB: 4B6U) is represented in cyan. TcEIF4E5-bound cap-4 and HsIF4E1-bound m⁷GTP are shown in sticks. The regions connecting β 1– β 2, β 3– β 4 and β 5– β 6 strands (S1–S2, S3–S4 and S5–S6) are indicated. (B, C and D) Detailed comparison between the m⁷GTP recognition sites. (B) TcEIF4E5-cap-4 (purple) versus HsIF4E1-m7GTP (green), (C) TcEIF4E5-cap-4 (purple) versus HsIF4E2-m7GTP (orange), (D) TcEIF4E5-cap-4 (purple) versus MmIF4E3-m7GTP (cyan). Cap-4 is represented in purple and m⁷GTP follows the color of the counterpart structure. The m⁷G group position is indicated. Residues involved in the interaction with cap analogues are shown in sticks and labeled, following the color of the protein structure.

in TcEIF4E5 has a short configuration in LmeIF4E1 similarly to the other eukaryotic counterparts.

Analysis of the TcEIF4E5 electrostatic potential surface evidences the positive charge mediating the interaction of the m⁷GTP phosphate groups while the remaining interacting surface is mostly neutral or lightly acidic (Figure 7A). This surface pattern most likely dictates the cap-4 interacting conformation, with the stacked bases towards the protein. In TcEIF4E5, we observe a cavity that accommodates the m⁷GTP formed by the regions connecting strands β 1–

β 2 and β 3– β 4, whereas the remaining cap-4 residues follow an open surface of the protein towards the long β 5– β 6 loop. Interestingly, the *L. major* EIF4E1 electrostatic potential surface strongly indicates a distinct cap-4 binding mechanism (Figure 7B). Four basic side chains present in the LmeIF4E1 S5–S6 loop (166-KRKYH-170) confer a positive potential to the corresponding surface. In contrast to TcEIF4E5, LmeIF4E1-cap-4 interaction would most likely occur with the phosphate groups of the AACU nucleotides oriented towards the protein, leaving the base moi-

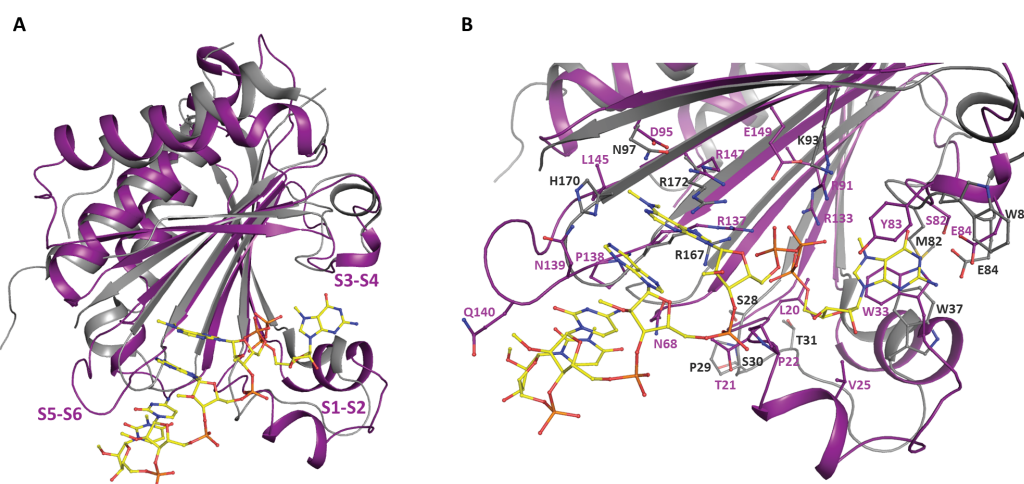


Figure 6. Structural comparison of *T. cruzi* EIF4E5-cap-4 complex and *L. major* EIF4E1 in the apo state. **(A)** Overall superposition: TcEIF4E5 is shown in purple and LmEIF4E1 (PDB 5WB5) is shown in gray. **(B)** Detailed comparison of the cap recognition sites. Cap-4 is represented in yellow. TcEIF4E5 residues involved in the interaction with cap-4 and corresponding residues in LmEIF4E1 are shown in sticks and labeled, following the color of the protein structure. The regions connecting $\beta 1$ - $\beta 2$, $\beta 3$ - $\beta 4$ and $\beta 5$ - $\beta 6$ strands (S1-S2, S3-S4 and S5-S6) are indicated.

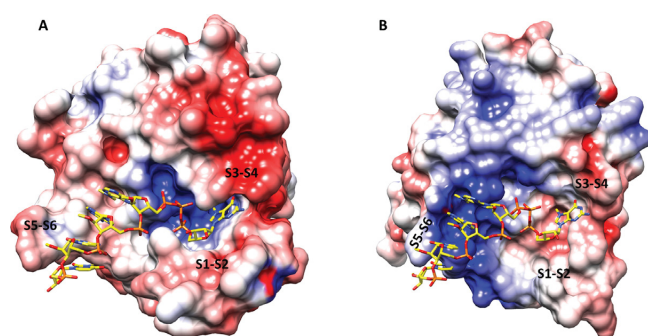


Figure 7. Electrostatic potential surface of **(A)** *T. cruzi* EIF4E5 and **(B)** *L. major* EIF4E1 (PDB 5WB5). The boundaries for potential contour map visualization are ± 5 kT/e. The cap-4 molecule (shown in sticks) is positioned on LmEIF4E1, based on the superposition of TcEIF4E5-cap-4 structure. The regions corresponding to the connections between $\beta 1$ - $\beta 2$, $\beta 3$ - $\beta 4$ and $\beta 5$ - $\beta 6$ strands (S1-S2, S3-S4 and S5-S6) are indicated. Electrostatic potential was calculated using the Adaptive Poisson-Boltzmann Solver through the software Chimera (45).

eties more exposed. Differences in EIF4E molecular mechanisms for cap binding can account for distinct mRNA affinities and are consistent with the proposal that trypanosomatid EIF4E homologues play distinct roles in translation.

T. cruzi EIF4E5 shows low affinity to cap analogues *in vitro* which is abolished by deletion of the C-terminal region

Affinity of TcEIF4E5 to m⁷GTP and cap-4 were investigated by microscale thermophoresis. Since the role of eIF4E C-terminal region in cap binding has been previously described in other organisms (4,5), a truncated TcEIF4E5 construction, lacking C-terminal residues 190 to 200 (TcEIF4E5- Δ C) was also used in the assays. Our results showed that TcEIF4E5 has low affinity to m⁷GTP and cap-4 ($K_d = 21 \pm 7$ and 110 ± 33 μ M, respectively) and interaction was abolished by deletion of the C-terminal region (Figure 8). TcEIF4E5 crystallographic models lack residues

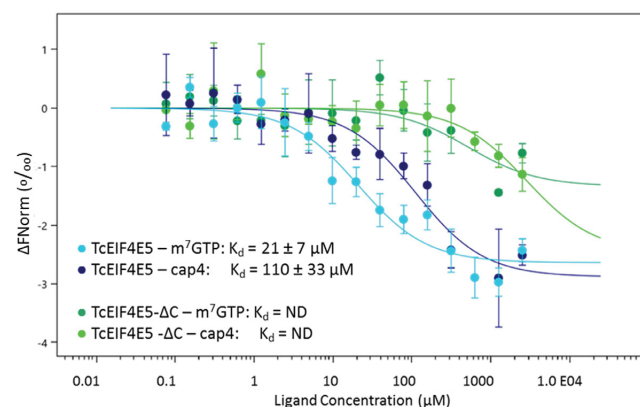


Figure 8. Microscale thermophoresis analysis of *T. cruzi* EIF4E5 binding to mRNA cap analogues. TcEIF4E5 presented low affinity to m⁷GTP and cap-4 (light and dark blue, $K_d = 21 \pm 7$ and 110 ± 33 μ M, respectively). Deletion of the C-terminal region abolishes the interaction (TcEIF4E5- Δ C, dark and light green).

187–200 even in the presence of cap-1 or cap-4. Therefore, although deletion of TcEIF4E5 residues 190–200 abolished cap binding, our structural analysis indicates that the role of C-terminal region is not in stabilizing the cap interaction.

The crystal structure of human eIF4E in complex with m⁷GpppA showed that the C-terminal loop is stabilized by interaction with the adenine nucleoside, while in the m⁷GTP complex the C-terminal remains disordered (4). In the crystal structure of parasitic *Schistosoma mansoni* eIF4E in complex with m⁷GpppA or m⁷GpppG, the C-terminal is also ordered, although a distinct orientation of the second base is observed, compared with the mammalian structure (5). In TcEIF4E5, cap-1 and cap-4 follow different orientations (Figure 9). Electron density was missing for cap-1 second base, indicating a flexibility for this nucleotide (Figure 3C). Cap-4 conformation is stabilized by the presence of the additional nucleotides, due to their inter-

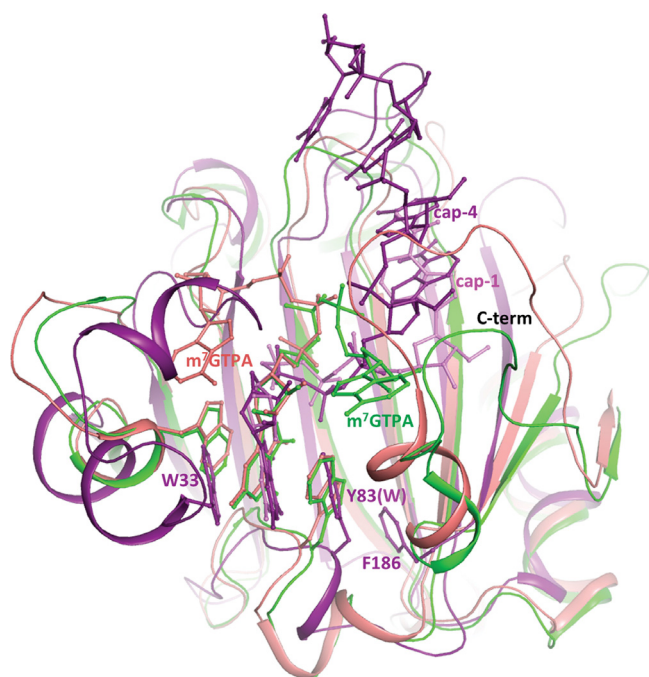


Figure 9. Structural superposition of TcEIF4E5-cap-4 (purple), human eIF4E1 (green) and *Schistosoma mansoni* eIF4E1 (salmon) in complex with m⁷GTPA (PDB: 1IPB and 3HXG, respectively). The ligands are represented in sticks and identified by the positions of the second base. Cap-1 (pink) is superimposed. The residues forming the m⁷G sandwich are represented in sticks and numbered according to TcEIF4E5 sequence. Human and *Schistosoma* eIF4E1 C-terminal regions, structured upon binding of a dinucleotide, are indicated. TcEIF4E5 model lacks C-terminal residues 187–200. The last residue built in TcEIF4E5 model (F186) is shown in sticks.

actions with the protein besides the stacking arrangement of the bases. Interestingly, four out of the eleven residues lacking in the TcEIF4E5-ΔC construction, and also absent from TcEIF4E5 crystal structure, are lysines. We hypothesize whether a flexible and basic C-terminal arm could function as a carrier, favoring the recruitment of the capped mRNA, before its positioning in the interaction pocket.

DISCUSSION

The interaction mode between EIF4E homologues and the mRNA cap in trypanosomatids is of special interest since the unique and extensive covalent modifications of cap-4 could be at the source of trypanosomatid-specific interaction mechanisms. The cap-4 structure includes multiple chemical modifications on the RNA body: the 2'-*O*-methylation on the first four ribose moieties and unique base methylations on the first adenine and the fourth uracil (m⁷Gpppm₃^{6,6,2'}Apm^{2'}Apm^{2'}Cpm₂^{3,2'}U) (compound 1, Figure 1). Access to mRNA cap structures remains a major bottleneck for structural studies of cap-binding proteins. Known synthetic routes to 5'-capped RNA are based on chemical or enzymatic preparation, as well as on a combination of both (46). A multi-step synthesis of trypanosome cap-4 has been previously reported on a 5–10 μmol scale by using RNA solid-support synthesis and phosphoramidite chemistry (30). However, a ten-fold excess of

phosphoramidite monomers is usually required to guarantee a high efficiency for each coupling step leading to the target oligomer in good overall yield. Scaling up the cap synthesis with this approach (i.e. solid support synthesis) would have required the preparation of rather large amounts of the corresponding phosphoramidite derivatives. In this work, these limitations were overcome by designing a new synthetic strategy providing large amounts of highly pure cap-4, suitable for structural and interaction analysis. Based on liquid-phase oligonucleotide synthesis (LPOS), it combines the beneficial aspects of conventional solid-phase and solution-phase chemistry (43). Our approach thus significantly reduced the amount of phosphoramidite building blocks required for RNA assembly, and consequently reduced the cost and time for the preparation of large quantity of cap-4.

T. cruzi EIF4E5 belongs to the so-called group 3 of kinetoplastid EIF4E homologues, which also includes EIF4E6. The functions of the extended EIF4E family in these organisms are poorly understood. The current knowledge, mostly obtained from studies in *T. brucei* and *L. major*, indicates that EIF4E3 and EIF4E4 form typical translation promoting eIF4F complexes and are the trypanosomatid homologues involved in general translation (reviewed in 20). *T. brucei* EIF4E5 is cytosolic and was shown to bind cap analogs *in vitro*. Knockdown of EIF4E5 in *T. brucei* resulted in loss of cell motility, suggesting a direct or indirect role in regulation of specific cellular pathways. In addition, the association of TbEIF4E5 with two distinct multiprotein complexes containing either the TbEIF4G1 or TbEIF4G2 homologues raised new questions about EIF4E5 function in trypanosomatids (25).

Structural comparisons of *T. cruzi* EIF4E5 with eIF4E family members achieved in this work pointed out significant differences in the cap recognition pocket. In addition to residue substitutions, long insertions in the TcEIF4E5 polypeptide chain, notably in the regions connecting β1-β2 and β5-β6 strands, are modifying the molecular interaction surface. Analysis of the electrostatic potential surfaces revealed a neutral to acidic surface, which is a unique feature of TcEIF4E5, providing the interaction path for the cap-4 AACU nucleotides (Figure 7A). Moreover, the cap-4 base moieties and methyl groups turned out to be involved in a number of hydrophobic interactions with the protein residues. In contrast, the eukaryotic counterparts, including *Leishmania* EIF4E1, feature an extended basic surface, that would be favorable to accommodate the ribonucleotide chain through an interaction with the phosphate groups (Figure 7B and Supplementary Figure S2). Structurally conserved waters were previously shown to participate in the cap binding interface. Fourteen water molecules are partially conserved in eIF4E structures in complex with cap analogues (47). In contrast, TcEIF4E5-m⁷GTP and TcEIF4E5-cap-4 structures present, respectively, seven and three waters in the cap pocket (Supplementary Figure S1). The reduced number of water molecules mediating TcEIF4E5-cap-4 interaction is a further distinct feature, which is most likely associated with its specific binding mechanism.

The differences in cap binding mode could be related to the role of the eIF4F-like complexes involving trypanoso-

matid homologues of EIF4E group 3. Sequence alignment of EIF4E5 and EIF4E6 homologues from *T. cruzi*, *T. brucei* and *L. major* (Supplementary Figure S3) pointed out a conservation of the extended loop connecting $\beta 5$ – $\beta 6$ strands, which is implicated, in TcEIF4E5, in cap-4 AACU interaction through the base moieties. However, the region connecting $\beta 1$ – $\beta 2$ strands, which includes one of the aromatic residues that participates in the m^7G stacking, is shorter in EIF4E6 when compared with EIF4E5 homologues (Supplementary Figure S3), indicating that, in EIF4E6, a distinct structural arrangement of the S1–S2 loop is likely to take place upon cap binding.

The crystal structures of *T. cruzi* EIF4E5 in complex with m^7GTP , cap-1 and cap-4 demonstrate the flexibility of the polypeptide chain in the extended region connecting $\beta 1$ – $\beta 2$ strands. Such flexibility was previously seen in the course of a structural comparison of human eIF4E2 in apo and m^7GTP bound states (14). The TcEIF4E5 S1–S2 region forms the base of the cavity which accommodates the m^7G nucleoside. Interestingly, electron density corresponding to S1–S2 residues is poor in the m^7GTP and cap-1 containing structures but is better defined in the cap-4 bound structure. The electron density map corresponding to the ligand is also of a better quality in the cap-4 containing structure (Figure 3). This density improvement indicates that the interactions involving N1–N4 nucleotides stabilize the cap-4 conformation and also the m^7GTP recognition pocket. In addition, our results suggest an indirect involvement of the flexible TcEIF4E5 C-terminal in the cap binding mechanism. We speculate whether the basic C-terminal arm could function as a carrier for the capped mRNA, through interaction with the phosphate groups.

Our biophysical analysis revealed a very low affinity of TcEIF4E5 for cap-4, which might be related to its specific interaction mode, engaging the cap-4 N1–N4 base moieties. Eukaryotic IF4E2 and IF4E3 were shown to bind the m^7GTP cap with affinities lower than the canonical eIF4E1 (14,15). The affinity of TcEIF4E5 for cap-4 is approximately 10-fold lower than the affinity of human eIF4E2 and mouse eIF4E3 for m^7GTP and 2000-fold lower than the affinity of canonical eIF4E1 for m^7GTP .

Studies on cap binding affinity of trypanosomatid EIF4E homologues have reported intriguing differences between homologues and species. Among *L. major* homologues, the highest binding affinity for m^7GTP and cap-4 was observed for EIF4E4 (equivalent dissociation constants of approximately 1.3 μM). *L. major* EIF4E1 showed preferential binding to methylated cap-4 (dissociation constants of 6.3 and 4 μM for m^7GTP and cap-4, respectively). On the other hand, a clear reduction in the binding affinity for cap-4 was observed for EIF4E3 (dissociation constants of 3.1 μM and 21.7 μM for m^7GTP and cap-4, respectively) (27). *T. brucei* EIF4E5 was reported to bind slightly better to m^7GTP than to cap-4 (dissociation constants of 1.5 and 1.8 μM , respectively) (25). Surprisingly, our interaction assays revealed the lowest affinity to cap analogues described for trypanosomatid homologues so far. We have confirmed our results using different batches of purified recombinant protein. Differences in binding constant values are expected when distinct experimental methods are applied. However, a difference of approximately two orders of magnitude as ob-

served for TcEIF4E5-cap-4 binding most likely indicates a very weak interaction, which is intrinsic to *T. cruzi* EIF4E5 homologue. Previous studies demonstrated that the interaction of eIF4E with the mRNA 5'-cap is enhanced by eIF4G (48). Moreover, in plants, association of PABP with the eIF4F complex was shown to increase the affinity of eIF4F for the cap structure (49). We hypothesize whether TcEIF4E5 affinity for cap-4 would be enhanced within the context of the translational machinery, or by the binding of an EIF4G factor. We also speculate about the influence of additional nucleotides of the spliced leader over the interaction of the mRNA with EIF4E in trypanosomatids. Further structural and interaction studies are necessary to elucidate the spectrum of molecular mechanisms for capped mRNA binding to the different trypanosomatid EIF4E homologues.

Biological data on *T. cruzi* EIF4E homologues are not available. Preliminary information on the interaction pattern between *T. cruzi* initiation factor homologues involved in the formation of eIF4F-like complexes indicates that TcEIF4E5 interacts with EIF4G1, EIF4G2 and EIF4G5 proteins (Dr. Fabiola Holetz, personal communication). TcEIF4E5 conserves the key residues that are implicated in the canonical 4E-4G and 4E-4EIP interactions, however the role of these interactions and the function of TcEIF4E5 in translation or other RNA metabolism pathway remain to be determined. The structural data presented in this work provide key starting points for the design of new experiments to investigate *T. cruzi* EIF4E5 function and cap-4 interaction mechanism with the remaining EIF4E homologues from trypanosomatids.

DATA AVAILABILITY

Atomic coordinates and structure factors for the reported crystal structures have been deposited with the Protein Data Bank under accession numbers 6O80, 6O7Z and 6O7Y for TcEIF4E5- m^7GTP , TcEIF4E5-cap-1 and TcEIF4E5-cap-4 complexes, respectively.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors acknowledge the Platform for Protein Purification and Characterization of the FIOCRUZ Technical Platform program. The authors are grateful to Synchrotron SOLEIL for beam time awarded at PROXIMA1 and PROXIMA2a beam lines (Proposal 20160778) and all the support they benefited as SOLEIL's users. We thank Jean-Sebastien Vartanian-Grimaldi for technical help during the preparation of ACSS. Yves L. Janin is also acknowledged for his help in proof-reading this manuscript. Electrostatic potential analyses were performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

FUNDING

Fundação Araucária [05/2016]; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES Financing code 001); Conselho Nacional de Desenvolvimento Científico e Tecnológico [CNPq, NITZ 312195/2015-0 BGG 304027/2015-4 and LWR fellowship]; Fundação Oswaldo Cruz. The authors also acknowledge financial support from Institut Pasteur and Centre National de la Recherche Scientifique (CNRS). Funding for open access charge: Fundação Oswaldo Cruz.

Conflict of interest statement. None declared.

REFERENCES

- Aitken, C.E. and Lorsch, J.R. (2012) A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.*, **19**, 568–576.
- Ramon, Y., Cajal, S., Castellvi, J., Hümmer, S., Peg, V., Pelletier, J. and Sonenberg, N. (2018) Beyond molecular tumor heterogeneity: protein synthesis takes control. *Oncogene*, **37**, 2490–2501.
- Marcotrigiano, J., Gingras, A.C., Sonenberg, N. and Burley, S.K. (1997) Cocystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell*, **89**, 951–961.
- Tomoo, K., Shen, X., Okabe, K., Nozoe, Y., Fukuhara, S., Morino, S., Ishida, T., Taniguchi, T., Hasegawa, H., Terashima, A. *et al.* (2002) Crystal structures of 7-methylguanosine 5'-triphosphate (m⁷GTP)- and P(1)-7-methylguanosine-P(3)-adenosine-5',5'-triphosphate (m⁷GpppA)-bound human full-length eukaryotic initiation factor 4E: biological importance of the C-terminal flexible region. *Biochem. J.*, **362**, 539–544.
- Liu, W., Zhao, R., McFarland, C., Kieft, J., Niedzwiecka, A., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E., Jones, D.N. and Davis, R.E. (2009) Structural insights into parasite eIF4E binding specificity for m⁷G and m^{2,2,7}G mRNA caps. *J. Biol. Chem.*, **284**, 31336–31349.
- Liu, W., Jankowska-Anyszka, M., Piecyk, K., Dickson, L., Wallace, A., Niedzwiecka, A., Stepinski, J., Stolarski, R., Darzynkiewicz, E., Kieft, J. *et al.* (2011) Structural basis for nematode eIF4E binding an m(2,2,7)G-Cap and its implications for translation initiation. *Nucleic Acids Res.*, **39**, 8820–8832.
- Marcotrigiano, J., Gingras, A.C., Sonenberg, N. and Burley, S.K. (1999) Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol. Cell*, **3**, 707–716.
- Gross, J.D., Moerke, N.J., von der Haar, T., Lugovskoy, A.A., Sachs, A.B., McCarthy, J.E. and Wagner, G. (2003) Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell*, **115**, 739–750.
- Umenaga, Y., Paku, K.S., In, Y., Ishida, T. and Tomoo, K. (2011) Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. *Biochem. Biophys. Res. Commun.*, **414**, 462–467.
- Grüner, S., Peter, D., Weber, R., Wohlbold, L., Chung, M.Y., Weichenrieder, O., Valkov, E., Igreja, C. and Izaurralde, E. (2016) The structures of eIF4E-eIF4G complexes reveal an extended interface to regulate translation initiation. *Mol. Cell*, **64**, 467–479.
- Grüner, S., Weber, R., Peter, D., Chung, M.Y., Igreja, C., Valkov, E. and Izaurralde, E. (2018) Structural motifs in eIF4G and 4E-BPs modulate their binding to eIF4E to regulate translation initiation in yeast. *Nucleic Acids Res.*, **46**, 6893–6908.
- Joshi, B., Cameron, A. and Jagus, R. (2004) Characterization of mammalian eIF4E-family members. *Eur. J. Biochem.*, **271**, 2189–2203.
- Joshi, B., Lee, K., Maeder, D.L. and Jagus, R. (2005) Phylogenetic analysis of eIF4E-family members. *BMC Evol. Biol.*, **5**, 48.
- Rosettani, P., Knapp, S., Vismara, M.-G., Rusconi, L. and Cameron, A.D. (2007) Structures of the human eIF4E homologous protein, h4EHP, in its m⁷GTP-bound and unliganded forms. *J. Mol. Biol.*, **368**, 691–705.
- Osborne, M.J., Volpon, L., Kornblatt, J.A., Culjkovic-Kraljic, B., Baguet, A. and Borden, K.L.B. (2013) eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 3877–3882.
- Rhoads, R.E. (2009) eIF4E: new family members, new binding partners, new roles. *J. Biol. Chem.*, **284**, 16711–16715.
- Hernández, G., Proud, C.G., Preiss, T. and Parsyan, A. (2012) On the Diversification of the Translation Apparatus across Eukaryotes. *Comp. Funct. Genomics*, **2012**, 256848.
- Bangs, J.D., Crain, P.F., Hashizume, T., McCloskey, J.A. and Boothroyd, J.C. (1992) Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *J. Biol. Chem.*, **267**, 9805–9815.
- Zinoviev, A. and Shapira, M. (2012) Evolutionary conservation and diversification of the translation initiation apparatus in trypanosomatids. *Comp. Funct. Genomics*, **2012**, 813718.
- Freire, E.R., Sturm, N.R., Campbell, D.A. and de Melo Neto, O.P. (2017) The role of cytoplasmic mRNA cap-binding protein complexes in *Trypanosoma brucei* and other trypanosomatids. *Pathogens*, **6**, E55.
- Jagus, R., Bachvaroff, T.R., Joshi, B. and Place, A.R. (2012) Diversity of eukaryotic translational initiation factor eIF4E in protists. *Comp. Funct. Genomics*, **2012**, 134839.
- Bannerman, B.P., Kramer, S., Dorrell, R.G. and Carrington, M. (2018) Multispecies reconstructions uncover widespread conservation, and lineage-specific elaborations in eukaryotic mRNA metabolism. *PLoS One*, **13**, e0192633.
- Dhalia, R., Reis, C.R., Freire, E.R., Rocha, P.O., Katz, R., Muniz, J.R., Standart, N. and de Melo Neto, O.P. (2005) Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. *Mol. Biochem. Parasitol.*, **140**, 23–41.
- Freire, E.R., Dhalia, R., Moura, D.M., da Costa Lima, T.D., Lima, R.P., Reis, C.R., Hughes, K., Figueiredo, R.C., Standart, N., Carrington, M. *et al.* (2011) The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. *Mol. Biochem. Parasitol.*, **176**, 25–36.
- Freire, E.R., Vashisht, A.A., Malvezzi, A.M., Zuberek, J., Langousis, G., Saada, E.A., Nascimento, J.F., Moura, D., Darzynkiewicz, E., Hill, K. *et al.* (2014) eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in *Trypanosoma brucei*. *RNA*, **20**, 1272–1286.
- Freire, E.R., Malvezzi, A.M., Vashisht, A.A., Zuberek, J., Saada, E.A., Langousis, G., Nascimento, J.D., Moura, D., Darzynkiewicz, E., Hill, K. *et al.* (2014) *Trypanosoma brucei* translation initiation factor homolog EIF4E6 forms a tripartite cytosolic complex with EIF4G5 and a capping enzyme homolog. *Eukaryot. Cell*, **13**, 896–908.
- Yoffe, Y., Zuberek, J., Lerer, A., Lewdorowicz, M., Stepinski, J., Altmann, M., Darzynkiewicz, E. and Shapira, M. (2006) Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in *Leishmania*. *Eukaryot. Cell*, **5**, 1969–1979.
- Meleppattu, S., Arthanari, H., Zinoviev, A., Boeszoermenyi, A., Wagner, G., Shapira, M. and Leger-Abraham, M. (2018) Structural basis for LeishIF4E-1 modulation by an interacting protein in the human parasite *Leishmania major*. *Nucleic Acids Res.*, **46**, 3791–3801.
- Yamamoto, I., Kimura, T., Tateoka, Y., Watanabe, K. and Ho, I.K. (1987) N-substituted oxypyrimidines and nucleosides: structure-activity relationship for hypnotic activity as central nervous system depressant. *J. Med. Chem.*, **30**, 2227–2231.
- Lewdorowicz, M., Yoffe, Y., Zuberek, J., Jemielity, J., Stepinski, J., Kierzek, R., Stolarski, R., Shapira, M. and Darzynkiewicz, E. (2004) Chemical synthesis and binding activity of the trypanosomatid cap-4 structure. *RNA*, **10**, 1469–1478.
- Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L. and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.*, **229**, 105–124.
- Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 125–132.
- Sheldrick, G.M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 479–485.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.*, **40**, 658–674.

35. Cowtan, K. (2010) Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 470–478.
36. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
37. Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.*, **62**, 1002–1011.
38. Vagin, A. and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 22–25.
39. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
40. Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonnrhein, C. and Womack, T.O. (2017) *BUSTER version 2.10.2*. Global Phasing Ltd., Cambridge.
41. Smart, O.S., Womack, T.O., Sharff, A., Flensburg, C., Keller, P., Paciorek, W., Vonnrhein, C. and Bricogne, G. (2011) *Grade, version 1.2.13*. Global Phasing Ltd., Cambridge.
42. Chen, V.B., Arendall, W.B. 3rd., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 12–21.
43. Kim, S., Matsumoto, M. and Chiba, K. (2013) Liquid-phase RNA synthesis by using alkyl-chain-soluble support. *Chem. Eur. J.*, **19**, 8615–8620.
44. Kadokura, M., Wada, T., Urashima, C. and Sekine, M. (1997) Efficient synthesis of γ -methyl-capped guanosine 5'-triphosphate as a 5'-terminal unique structure of U6 RNA via a new triphosphate bond formation involving activation of methyl phosphorimidazolidate using ZnCl_2 as a catalyst in DMF under anhydrous conditions. *Tetrahedron Lett.*, **38**, 8359–8362.
45. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
46. Muttach, F., Muthmann, N. and Rentmeister, A. (2013) Synthetic mRNA capping. *Beilstein J. Org. Chem.*, **13**, 2819–2832.
47. Lama, D., Pradhan, M.R., Brown, C.J., Eapen, R.S., Joseph, T.L., Kwok, C.K., Lane, D.P. and Verma, C.S. (2017) Water-bridge mediates recognition of mRNA cap in eIF4E. *Structure*, **25**, 188–194.
48. Haghighat, A. and Sonenberg, N. (1997) eIF4G dramatically enhances the binding of eIF4E to the mRNA 5'-cap structure. *J. Biol. Chem.*, **272**, 21677–21680.
49. Wei, C.C., Balasta, M.L., Ren, J. and Goss, D.J. (1998) Wheat germ poly(A) binding protein enhances the binding affinity of eukaryotic initiation factor 4F and (iso)4F for cap analogues. *Biochemistry*, **37**, 1910–1916.